DOCUMENT RESUME

ED 458 887 IR 058 387

AUTHOR	Nichols, Stephen G.; Smith, Abby
TITLE	The Evidence in Hand: Report of the Task Force on the Artifact in Library Collections. Optimizing Collections and Services for Scholarly Use.
INSTITUTION	Council on Library and Information Resources, Washington, DC.
ISBN	ISBN-1-887334-88-2
PUB DATE	2001-11-00
NOTE	125p.
AVAILABLE FROM	Council on Library and Information Resources, 1755 Massachusetts Ave., NW, Suite 500, Washington, DC 20036 ($20). Web site: http://www.clir.org. For full text: http://www.clir.org/pubs/reports/pub103/contents.html.
PUB TYPE	Information Analyses (070)
EDRS PRICE	MF01/PC05 Plus Postage.
DESCRIPTORS	Access to Information; Audiovisual Aids; Higher Education; Information Technology; *Library Materials; *Library Services; *Preservation; Printed Materials; Research Libraries
IDENTIFIERS	*Artifacts; Digital Data; Digitizing

ABSTRACT

This report discusses the ways in which artifacts and their physical frailties affect their research value--originality, faithfulness, fixity, and stability--over time and how libraries can minimize the risk of unacceptable loss of that value. It investigates the specific issues around selected media, such as paper-based printed matter, moving image and recorded sound materials, and objects that exist in digital form, and points to two key strategies currently in use in libraries that can be scaled up to deal with a problem of this magnitude. The first strategy is preventive maintenance, which most often entails storing items under optimal conditions to retard rates of natural decay. Prevention also includes such techniques as deacidifying items printed on wood-pulp paper, stabilizing bindings and encasing fragile items, and improving care and handling techniques for all library materials. The second strategy is the use of surrogates to reduce the stress of handling. For research purposes, digital surrogates in particular were found to be quite acceptable and were even preferable to the originals in a number of cases. Surrogates, especially when networked, have the added benefit of increasing access to an item and of providing convenience of access to items housed at distant or disparate locations. (Contains 68 references.) (MES)

Reproductions supplied by EDRS are the best that can be made
from the original document.

ED 458 887

The Evidence in Hand:

Report of the Task Force on the Artifact in Library Collections

November 2001

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
This document has been reproduced as received from the person or organization originating it.
Minor changes have been made to improve reproduction quality.
Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY
B.H. Leneyr
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
1

BEST COPY AVAILABLE

Council on Library and Information Resources

2

This report is one in a series from CLIR on
OPTIMIZING COLLECTIONS AND SERVICES FOR SCHOLARLY USE

IR058387

The Evidence in Hand: Report of the Task Force on the Artifact in Library Collections

Stephen G. Nichols, Chairman
Task Force on the Artifact in Library Collections
Johns Hopkins University

Abby Smith
Director of Programs
Council on Library and Information Resources

November 2001

Council on Library and Information Resources
Washington, D.C.

Cover image: "Scene in the Old Congressional Library, Washington, D.C., Showing Present Congested Condition," by W. Bengough. The illustration appeared in *Harper's Weekly* on February 27, 1897. CLIR is grateful to The Making of America Digital Collection, Cornell University Library, for providing the scanned image.

ISBN 1-887334-88-2

Published by:

Council on Library and Information Resources
1755 Massachusetts Avenue, NW, Suite 500
Washington, DC 20036

Web site at http://www.clir.org

Additional copies are available for $20 per copy. Orders must be placed online through CLIR's Web site.

The paper in this publication meets the minimum requirements of the American National Standard for Information Sciences—Permanence of Paper for Printed Library Materials ANSI Z39.48-1984.

Copyright 2001 by the Council on Library and Information Resources. No part of this publication may be reproduced or transcribed in any form without permission of the publisher. Requests for reproduction should be submitted to the Director of Communications at the Council on Library and Information Resources.

Contents

EXECUTIVE SUMMARY v

TASK FORCE MEMBERS ix

1. THE PROBLEM 1

    1.1 The Charge 5
    1.2 The Work of the Task Force 7

2. THE ARTIFACT IN QUESTION 8

    2.1 Selection for Preservation of the Original 9
    2.2 Frameworks for Determining Value 12

3. STATES OF THE ARTIFACT, 1800–2000 17

    3.1 The Changing Features of the Artifact 17

    3.2 Print/Paper 19
        3.2.1 The Relative Stability of Imprints 19
        3.2.2 Evaluation of the Artifact and Selection for Preservation 23
        3.2.3 Creating Surrogates: Filming versus Scanning 24
        3.2.4 Responsible Retention Policies 28

    3.3 Audiovisual 30
        3.3.1 Sound and Light as Artifacts 30
        3.3.2 Still Images 33
        3.3.3 Moving Images 35
        3.3.4 Recorded Sound 38
        3.3.5 Broadcast Media 40

    3.4 Digital 41
        3.4.1 Artifacts and Artifactual Value in the Digital Realm 42
        3.4.2 Digital Surrogates 43
        3.4.3 Access 46
        3.4.4 Born-Digital Materials 47
        3.4.5 Preservation of Digital Information 50
        3.4.6 Copyright: A Barrier to Preservation? 53

4. CASE STUDIES 55

4.1 The Five-College Library Depository 55
4.2 *The Emperor Jones* 58
4.3 Preserving Oral Traditions 61
4.4 JSTOR 64
4.5 The Rossetti Archive 66

5. SUMMARY AND RECOMMENDATIONS 70

5.1 Principles of Good Stewardship 72
5.2 Best Practices for Preservation of the Artifact 72
5.3 Strategies for Specific Formats 72
5.4 Recommendations 73
5.5 Areas for Further Research 74

6. REFERENCES 75

7. APPENDIXES 82

I. Current Library Practices in Collection Development and Preservation ... 83
II. Charge to the Task Force 91
III. Selection for Preservation Criteria from RLG, Cornell, and Harvard 93
IV. Expenditure Trends in ARL Libraries, 1986–2000 99
V. Non-Print Holdings in ARL University Libraries 100
VI. Comparative Costs for Book Treatments 101
VII. National Recording Preservation Act of 2000 104

Executive Summary

The "information explosion" sparked by digital technology has fostered an increasing awareness of the sheer mass of information available today in a variety of media, from traditional formats such as paper to the more recent film, optical, and magnetic formats. Institutions charged with collecting, storing, preserving, and making accessible recorded information are struggling to keep pace with the growth of information production, even though their brief is to collect only a portion of what is published and an even smaller portion of what is produced and disseminated in unpublished form. With so much information produced, how do members of the research community—scholars and teachers, librarians and archivists, and academic officers who support their work—distinguish between what is of long-term value, what is ephemeral, and what of that ephemera is valuable for the preservation of a rich historical record?

Libraries have struggled with these questions for a long time, but the issue of how to understand and protect the value in research collections, in particular in the original, unreformatted materials in physical formats that abound in libraries and archives, is more urgent than ever. The Council on Library and Information Resources convened a task force of scholars, academic officers, librarians, and archivists to investigate the role of artifacts—original, unreformatted materials—in libraries and archives as these institutions are creating digital collections and services to serve research and teaching functions. The task force members were asked to articulate a general context or framework for formulating and evaluating institutional policies on how best to preserve and make accessible artifactual collections and how to ensure their continued access for research and teaching needs, however those needs evolve.

The task force was charged to answer specifically these questions:

- What qualities of an original are useful or necessary to retain in their original form? Under what circumstances are original materials required for research?
- When is it sufficient and appropriate to capture intellectual content through reformatting and not necessarily retain the original?
- Which preservation options provide the most appropriate and cost-effective means of preserving the original?
- From both custodial and scholarly perspectives, what are the advantages and disadvantages of these various preservation options?

Task force members addressed the needs of print and paper-based collections; of audiovisual materials, including still and moving images, recorded sound, and broadcast media; and of digital formats, both those created from analog sources and those that are created and exist exclusively in digital form.

The challenges to preserving artifactual collections are formidable. Artifacts are at risk because they are produced in *high quantity*; because they are recorded onto *unstable media* that decay over time; and because the *economics of preservation* result in preservation needs of collections competing less and less successfully with the access needs of users. Moreover, developing priorities for preservation actions on the basis of the research value of an object is challenging because of the *unknown and unfixed values of artifacts*. The dynamic nature of intellectual inquiry increases demand for some types of sources while neglecting others, and competition among communities of interest can result in difficult trade-offs.

Despite these challenges, a number of critical research functions will continue to depend on access to the original. The task force found that preservation budgets in research libraries have been flat—in real terms declining—since 1994, and that they often fail to meet the preservation needs of artifacts. The attention given to increasing access through digital reformatting has the potential to eclipse the preservation needs of artifacts and to preoccupy the attention of the research community. The needs of audiovisual materials seem particularly serious, given their sheer number, the fact that relatively few institutions are collecting them in original form, and the fact that preservation issues relating to these media are difficult and expensive to address. Scholars' general lack of awareness of the value and endangered nature of these materials leaves them with few advocates for preservation resources.

The task force looked at traditional criteria used in selecting for preservation—age, rarity, associational value, evidentiary value—and found these criteria still valid. Because artifacts of evidentiary value often have little market value and are found in multiple copies, decisions about their treatment and retention are often contested; for this reason, the task force focused chiefly on these items. Artifactual collections that are paper-based or audiovisual have evidentiary value to the extent that the original manifestation can attest to the *originality*, *faithfulness* (or authenticity), *fixity*, and *stability* of the content. When the items are reformatted onto newer media for preservation or other purposes, those qualities inalienable to the original may be lost. When that is the case, the artifact should be retained.

The report deals in detail with the ways in which artifacts and their physical frailties affect their research value—originality, faithfulness, fixity, and stability—over time and with how libraries can minimize the risk of unacceptable loss of that value. It investigates the specific issues around selected media, such as paper-based printed matter, moving image and recorded sound materials, and objects that exist in digital form, and points to two key strategies currently in use in libraries that can be scaled up to deal with a problem of this magnitude.

The first strategy is *preventive maintenance*, which most often entails storing items under optimal conditions to retard rates of natural decay. Prevention also includes such techniques as deacidifying items printed on wood-pulp paper, stabilizing bindings and encasing fragile items, and improving care and handling techniques for all library materials.

The second strategy is the *use of surrogates* to reduce the stress of handling. For research purposes, digital surrogates in particular were found to be quite acceptable, and were even preferable to the originals in a number of cases. Surrogates, especially when networked, have the added benefit of increas-

ing access to an item and of providing convenience of access to items housed at distant or disparate locations.

The task force acknowledges that both strategies have their drawbacks. For preventive measures, the chief drawback may be that optimal storage conditions are created only in closed-stack environments. This practice has great benefit to the collections but can affect the ways in which researchers gain access to them. With the use of surrogates, the drawbacks from the user's point of view depend chiefly on the type of surrogate created, with digital scans and preservation photocopying being preferred to microforms. The larger problem may be that, as preference for the convenience of surrogates grows in the digital library, the continuing preservation needs of the source artifacts may be eclipsed by the resource needs of creating and maintaining digital files, the long-term costs of which are unknown.

Given the size of the collections that are of potential research value over time, the desirability of ensuring that a meaningful number of them are preserved for present and future access purposes, and the desirability of maintaining the richest possible historical base of cultural and intellectual resources, the task force concluded that librarians and scholars need to build economies of scale into preservation strategies.

Specific recommendations for these economies of scale for different media follow.

Print/Paper

- Establish regional repositories to house and provide proper treatment of low-use print matter. These repositories would provide access to artifacts aggregated from different institutions under terms to be worked out. Such repositories might begin by taking in journals and monographs that are widely available in digital form.
- Convene a national group to investigate the establishment of archival repositories that would retain a "last, best copy" of American imprints.
- Build interinstitutional networks of information sharing about the status of artifacts and delegation of responsibilities for caring for them. These networks would obviate unwanted duplication and encourage libraries to take responsibility for the preservation of their most important or rarely held materials. These networks would include a registry of digitized collections that has information about where the originals are located, who has responsibility for them, and under what terms they may be used.

Audiovisual

- Extend the reach of the National Film Preservation Plan sponsored by the National Film Preservation Board to continue preserving historically significant films and, in particular, to serve the needs of noncommercial films.
- Implement and extend the reach of the national plan to preserve television and video that has been proposed in a study by the Library of Congress. The U.S. Congress has authorized a similar assessment and plan for recorded sound.
- Encourage scholars and librarians to identify important media collections in need of preservation and set priorities for treatment.
- Urge scholars and librarians to develop controlled vocabularies and common descriptive standards for the intellectual and inventory control of media collections.

Digital

- Develop sound criteria for selecting collections to digitize.
- Develop benchmarks and minimal-level standards for capture and metadata, and document the specifications used.
- Create registries of digitized artifacts to document what has been converted and how, and to facilitate access to the source materials as well as to the surrogates.
- Develop and use nonproprietary formats for creating born-digital objects to facilitate the creation of preservable digital files.
- Urge scholars and librarians to develop criteria for selecting born-digital materials to be preserved or managed for long-term access.
- Continue to research and develop methods for preserving digital files.

Task force members were asked to look only at those materials around which a debate could arise on retention of the original. They recognized that the preservation needs of special collections and rare-book collections are also great and should be addressed in any preservation strategy at both the local and the national levels.

Good stewardship of the intellectual and cultural assets in libraries and archives is the responsibility of all members of the research community, not merely of the librarians and archivists who have immediate custody of the collections. The preservation challenge cannot be deferred or deflected, for what is lost by the present generation cannot be retrieved by the next.

Stephen G. Nichols
Chairman, Task Force on the Artifact in Library Collections
Johns Hopkins University

Abby Smith
Director of Programs
Council on Library and Information Resources

Task Force Members

Stephen G. Nichols, Chairman
James M. Beall Professor of French and Humanities
Chair, Romance Languages
Johns Hopkins University

Francis X. Blouin
Director
Bentley Historical Library
The University of Michigan

Bernard Cerquiglini
Director
Institut national de la langue française

Rebecca S. Chopp
Provost and Vice President for Academic Affairs
Emory University

Sheldon Hackney
Professor of United States History
University of Pennsylvania

Charles Méla
Université de Genève

James J. O'Donnell
Vice Provost, Information Systems/Computing
University of Pennsylvania

Henry Petroski
A. S. Vesic Professor
Department of Civil Engineering
Duke University

Abby Smith
Director of Programs
Council on Library and Information Resources

Sarah Thomas
University Librarian
Cornell University Library

John Unsworth
Director
Institute for Advanced Technology in the Humanities
University of Virginia

Nancy Vickers
President
Bryn Mawr College

Steven C. Wheatley
Vice President
American Council of Learned Societies

Karin Wittenborg
University Librarian
University of Virginia

Pauline Yu
Dean of Humanities
College of Letters and Science
University of California at Los Angeles

1. The Problem

The "information explosion" sparked by digital technology has fostered an increasing awareness of the sheer mass of information available today in a variety of media, from traditional formats such as paper to the more recent film, optical, and magnetic formats.[1] Institutions charged with collecting, storing, preserving, and making accessible recorded information are struggling to keep pace with the growth of information production, even though their brief is to collect only a portion of what is published and an even smaller portion of what is produced and disseminated in unpublished form. With so much information coming at us, how do we distinguish between what is of long-term value and what is ephemeral? And of that ephemera, what should be selected to preserve to ensure for the future a rich record of the present?

Libraries have struggled with these questions for a long time. Even aggressive collecting will not meet all present and future information needs. In 1875, the Librarian of Congress noted that the quarters the library occupied in the U.S. Capitol were inadequate for the burgeoning collections and, in a thinly veiled appeal to Congress for a new building, described the problems the library had with newspapers and journals.

> Though carefully preserved and promptly bound for preservation, there is no longer the possibility of even receiving half the issues of these representative journals, so important in our current history and politics; and the time will soon come when a legislator in search of a fact, a date, a political article, or a table of statistics known to be in a certain newspaper at a certain

[1] One study reports that the world produces a startling 2 billion gigabytes of new information a year, or roughly 250 megabytes for every man, woman, and child on earth (Lyman and Varian 2000).

> date, will find it only at the bottom of a lofty pile of journals, all of which must be displaced before it can be reached. Besides the issues of the daily press, the periodicals which are taken under the copyright law or by subscription, embracing most of the monthly and quarterly magazines and reviews, accumulate with such rapidity that no device yet invented will long avail to produce them when wanted. (Cox 2000)

Information is of little long-term use if it is not preserved and made accessible through indexing or cataloging. At the same time, much of what has been collected and made accessible by libraries and archives admittedly receives little or no use. This leads some to argue that investments in preserving so much material that has so little demonstrated use are a waste of resources. This point of view is nothing new. In the 1890s, around the time when the Library of Congress was moving its collections into a new, purpose-built building, the consolidation of the Lenox and Astor Libraries and Tilden Foundation in New York to create the New York Public Library led to the collections being closely scrutinized. This move prompted the following letter to the *New York Herald* from a concerned citizen.

> To the Editor of the *Herald*:
>
> Before the Astor Library moves its quarters, it would do well to get rid of some scores of thousands of its untouched volumes in order to make room for readable books. When, after asking in vain for some work that is the talk of the town, I look up and around at the rows of shelves packed with "things in books' clothing," I find myself echoing the poet's wish that
>
> > "... From the dead
> > Old Omar would pop forth his head,
> > And make a bonfire of them all."
>
> Like Charles Lamb, I can read anything that I call a book, but there is much in the Astor Library which, in my humble opinion, should be categoried as waste paper.
>
> Clara Marshall
> New York, April 6, 1895

Today, little has changed. There is ongoing tension between how much information is produced and how much can be acquired, preserved, and made accessible in meaningful ways. There is tension between those who think we should collect as broadly as possible to expand our research base and those who think too much information can impede one's ability to find meaningful information. And there is always the question of who should be bearing the burden of whatever preservation society deems necessary.

Currently, the Library of Congress has more than 17 million books and 95 million manuscripts, films, photographs, maps, sound recordings, and other non-book items. The New York Public Library

has 42 million items, including more than 14 million books. Despite these impressive numbers, these libraries hold only a portion of what is published, or of what is created but not published. The Library of Congress, for example, which receives items that are registered for U.S. copyright deposit as well as foreign materials, reports that approximately 31,000 items arrive at its doorstep each working day. Of these, about 7,000 are selected to become part of its permanent collections (LC 2001). If libraries, archives, or historical societies do not collect instances of recorded information, then the chance of their survival is slim. Loss is inevitable.

Although information overload is not a new problem, the introduction of digital technology onto campuses and into research libraries has fundamentally altered the information landscape and created problems for scholars and students that have potentially serious ramifications. The creation and dissemination of digital resources are creating new models of service and access, such as licensing rather than owning essential intellectual assets. The mutability of digital documents is redefining what constitutes a text. For example, are back issues of a journal that are in digital form simply a bunch of articles or a rich database? Moreover, the conversion of texts into searchable texts is resulting in increased interdisciplinary research, as researchers in one field serendipitously find resources that had heretofore been confined (in print) to another field.

Accompanying the trend toward networked information are others that at first seem paradoxical. For example, at the very time that more material is made available online and retrievable any time, anywhere, there is increased attention among scholars to original, unreformatted materials and an increased appreciation for the material aspects of these sources (Tanselle 1998). There are eruptions of public outcry when material losses in libraries and archives are discovered. Scholars demand increasing attention to an ever-expanding range of candidates for preservation, but library budgets simply cannot support those demands. Preservation has thus become an unfunded mandate, the more pernicious for often being implicit. Academic institutions have learned the huge costs of penny-wise facilities management and deferred maintenance. It is reasonable to fear that libraries are incurring future costs by deferring preservation.

There is, however, a crucial difference between deferring maintenance costs and deferring expenses for preservation. When we do not allocate sufficient funds for preservation, we face the probability that significant resources will be lost forever. Library collections are among the most valued of a research institution's intellectual and cultural assets—assets that form a crucial part of what might be called "public goods." Unfortunately, with library collections, as with other public goods, many of those who make claim to their use are not prepared to take responsibility for their well-being.

This report addresses the value of library collections, especially those in physical formats, to scholarship and teaching. It discusses the costs of the benefits these collections bring and the roles of each

member of the various communities who have a claim on their use and a responsibility for their well-being. The report focuses on preservation—what it takes to ensure the present and future usability of collections. It is grounded in the recognition that without preservation today, there will be no access tomorrow.

Preservation is a critical part of good stewardship of our intellectual and cultural heritage. Its chief challenges at the turn of the twenty-first century are fourfold:

Quantity. Because of the relentless growth of research libraries and their collections, an immense number of research items demand resources to remain accessible. In 1999, the 121 member libraries of the Association of Research Libraries (ARL) reported owning a total of 462,965,204 volumes (ARL 2000). The three greatest periods of growth for research libraries occurred after World War I, after World War II, and during the 1990s. A look at 12 representative public university libraries during this period reveals a typical growth pattern.[2] In 1907, these libraries held an average of 107,425 volumes. By 1961, that number had grown to 1,772,831, and by 1995, the average number of volumes per library was 5,334,620 (Molyneux 1998).

Stability of Media. Library collections exist in a variety of physical formats, all of which are vulnerable to some degree. As the rate of information production has increased, storage media have become more compact and efficient. However, miniaturization comes at the expense of stability and longevity. With the exception of preservation-quality microfilm, the new media of the twentieth century are more fragile than those of the nineteenth, including the infamous wood-pulp paper that has been known to deteriorate into flakes over time (Conway 1996). The media invented in the last 150 years to capture light and sound are generally extremely fragile, dependent on machines for playback, and subject to rapid technological obsolescence. For example, the wax cylinders on which are inscribed the earliest known recorded voices of Native Americans are susceptible to mold, heat, scratching, skin oils, and other physical trauma. Moreover, they are dependent on playback equipment for which no replacement parts are manufactured and that few technicians are able to repair. Nonetheless, the information on them is invaluable and should be saved for future generations.

Economics. Since 1993, preservation budgets in ARL libraries have remained flat. The number of staff assigned to preservation is at a 10-year low (Reed-Scott 1999). At the same time, the demand for access to original materials has grown, especially for access to those materials in special formats that often are at greatest risk from physical handling or environmental stress. Meanwhile, although the technology for reformatting for access has greatly improved, the funding for preservation continues to decrease. More money now goes to dig-

[2] The libraries are University of California, Berkeley; University of Illinois, Urbana-Champaign; Indiana University; University of Iowa; University of Kansas; University of Michigan; University of Minnesota; University of Missouri; University of Nebraska; Ohio State University; University of Washington, Seattle; and the University of Wisconsin, Madison.

ital reformatting of items to provide access than to microfilming to preserve the low-use brittle books that are rotting on shelves.

Unknown and Unfixed Values of Artifacts. The most difficult challenge for libraries is deciding how to set priorities for preservation. As long as the claim on preservation resources exceeds the available funds, it will be necessary to select which materials will get treatment and which will not. The choices are made in the context of changing perceptions of value and the fluid dynamics of intellectual inquiry. The recent elevation of nineteenth-century popular imprints and ephemera to a status of high research value poses an exemplary challenge. Providing access to those items, which are often at high risk from embrittlement and routine physical handling, has put great strain on library resources. Knowing that the intellectual interests and research methodologies of scholars will change over time, research institutions have collected "just in case" there is a demand in the future, rather than "just in time" for current demand. Research libraries and archives are full of items that have not been consulted in decades, if ever, and for which future demand is unpredictable, yet which make their claim for preservation attention.

How do a library and its home institution make sound fiscal and intellectual decisions about what to preserve, when, for whom, and at what price? Despite the enormous collections of printed materials that have been amassed, entire categories of primary sources have disappeared before collecting institutions and their users understood their value. A notorious example of such neglect is the fact that 80 percent of all silent films made in the United States are gone without a trace. Fifty percent of films made in the nitrate era (that is, before 1950) have also perished. Among those extant, a significant portion are not well preserved. Given that the materials that have vanished were not well documented at the time of their creation, the full extent of this loss will never be known.

1.1 The Charge

To seek several perspectives on the importance of the artifact to academic inquiry and teaching and to propose strategies for addressing the problems just defined, the Council on Library and Information Resources (CLIR) created the Task Force on the Artifact in Library Collections in 1999. The task force was charged with answering a challenging question: What will constitute good stewardship of our intellectual and cultural heritage in the first decades of the twenty-first century?

Scholars and librarians have common interests in the collection-building process. Although scholars are routinely, and often intimately, engaged with librarians in the acquisitions and use phases of collection building, they are usually exposed only to specific aspects of preservation activities, such as testifying to the value of a collection proposed for preservation microfilming. Rarely do scholars have the opportunity to view the preservation function as a whole.

Recent vigorous debates in the scholarly community and in the public arena about the value of saving books and newspapers as artifacts, in addition to preserving their content through reformatting and extending access through digitization, have raised questions about the involvement of scholars in these decisions. Do scholars know what happens to original materials after preservation treatment? Do they participate in decision making about the disposition of original materials after preservation treatment or reformatting? Are provosts and other academic officers fully informed about the economic factors that may constrain a library's ability to keep its collections fit for use? The task force was asked to engage in a systematic review of the scholarly community's interests in this and to explore the broader question of how to develop a vision for building collections in a new information environment comprising both physical and digital materials.

The task force was asked to articulate a framework for making or evaluating institutional policies for the retention of published materials and archival or unpublished materials in their original form. The crucial questions associated with this task were as follows:

- What qualities of an original are useful or necessary to retain in their original form? Under what circumstances are original materials required for research?
- When is it sufficient and appropriate to capture intellectual content through reformatting and not necessarily to retain the original?
- Which preservation options provide the most appropriate and cost-effective means of preserving the original?
- From both custodial and scholarly perspectives, what are the advantages and disadvantages of these various preservation options?

Given that resources for preservation are finite and duplication of efforts can be costly, the task force was also asked to consider the advisability and feasibility of creating one or more national repositories into which one or more copies of all materials published in the United States would be deposited and permanently retained.

The task force was asked to interpret its charge broadly. It was asked to give primary consideration to print formats but also to consider the burgeoning numbers of non-print and electronic research sources that increasingly demand attention from preservation specialists. In fact, shortly after the task force convened, its members identified the problems facing non-print materials as being as important, certainly larger in scope, and probably more urgent than those facing print materials. The task force therefore gave almost equal consideration to three types of collection materials: print, analog audiovisual, and digital. Although technical constraints and considerations of preservation actions cannot be ignored, members were urged to focus on their implications for the research process, rather than on the technologies themselves. Preservation and access are technology-dependent, and best practices for ensuring long-term access to information will change over time.

1.2 The Work of the Task Force

In taking up its charge, the task force confirmed that the pressure on research libraries to keep up with the past while preparing for the future is of vital interest to scholars, governing boards, academic officers, and funders, as well as to librarians and archivists. Members agreed on several premises that would guide the task force's work. Although these premises can be simply stated, each speaks to the uncertainty and dynamism that characterize the environment in which research libraries operate. The premises are as follows:

- Information technology will continue to change rapidly.
- Best practices for preservation and access will change.
- Digital resources will increase significantly.
- Scholarly research trends will change; in anticipation of such changes, scholars will continue to demand that collections be as inclusive as possible.
- Intellectual property-rights management will evolve and must be respected.
- Financial and human resources will not keep pace with demand.
- Financial and human resources should be allocated in the most cost-effective manner to achieve an acceptable trade-off between quality of resource and expenditure of time and money.

Task force members began with a survey of current library practices in preservation and in collection development. They gathered information about what research libraries do and do not collect, how libraries will ensure access to those collections over time, and how they set priorities for investments in their collections, especially for preservation. (These findings are summarized in Appendix I.) Task force members then identified the particular needs of library materials in three areas: print, analog audiovisual, and digital. They determined that the ways in which these media are collected and preserved—and even which institutions do or do not engage in certain activities—are distinct issues that warrant separate consideration (see Section 3, States of the Artifact, 1800–2000). Throughout their inquiry, task force members returned to the simple but almost overwhelming fact that the web of resources that are useful for scholarship continues to grow at a rapid pace. More information is being produced and disseminated than before, and ever-widening orbits of intellectual inquiry call for ready access to these materials. While many of these sources, from business records to digital art, fall outside the traditional purview of library collecting, they are part of the larger information environment that influence public expectations of what libraries can and should be doing.

The task force relied on the testimony of experts from the library and archival fields for information ranging from the technical to the financial to the theoretical. On the basis of these findings and in consideration of the economic realities of funding preservation and access, the task force proposed definitions of responsible stewardship of research and cultural materials, and articulated the roles that each party in the research community and beyond plays in that stewardship (see Section 5, Summary and Recommendations).

BEST COPY AVAILABLE

2. The Artifact In Question

The word "artifact" can be confusing because it masks a number of unexamined assumptions. In academic parlance, "artifact" can refer to a physical object, a primary record, or a physical object that constitutes a primary record.[3] For the purposes of this report, an artifact will be defined as an information resource in which the information is recorded on a physical medium, such as a photograph or a book, and in which the information value of the resource adheres not only in the text or content but also in the object itself. In other words, artifacts are things that have intrinsic value as objects, independent of their informational content.

In recent years, scholars have identified an increasing number of library items that have research value as physical objects, above and beyond the information recorded in them. The Modern Language Association (MLA) has been concerned that the technologies of reproduction, such as photocopying, microfilming, and digital scanning, are becoming so good, so readily available, and so serviceable for many research and teaching purposes, that the importance of the underlying original might be devalued. To call attention to the dangers inherent in ignoring the fate of physical collections, the association created a committee to consider the issue. This group defined an artifact or primary record as "a physical object produced or used at the particular past time that one is concerned with in a given instance" (MLA 1996). Members asserted that for practical purposes, all historical publications, even those produced by mass-production techniques designed to minimize deviations from a norm, have unique physical qualities that may have value as a carrier of (physical) evidence in a given research project.[4] Although careful to note it

[3] In scientific laboratories, "artifact" also denotes a phenomenon or substance that is a by-product of some external action or agent.

[4] The statement addresses only text-based sources. If this standard of value were extended to visual and sound resources, the universe of primary records would grow exponentially.

is not possible to save all copies of printed materials from destruction and the ravages of time, the committee's statement nonetheless provoked some anxiety among librarians. This is because, while asserting the importance of preserving as many artifacts as possible and acknowledging the need to set priorities for preservation, the statement gave no guidance about how to make such priorities. Nicholson Baker recently alleged that libraries' poor stewardship of books and serials has resulted in the loss of many resources of artifactual value. He further asserted that it is the responsibility of libraries in general and of certain large libraries in particular to collect masses of primary source materials and preserve them in their artifactual form. Yet he, too, failed to address the crucial matters of who would bear the responsibility for setting priorities, who would assume the custodial burden of these comprehensive collections, and who would fund these activities (Baker 2001).

An increasing number of library collections are being promoted, as it were, to resources of artifactual, not just informational, value. Given the nature of contemporary scholarship and its wide-ranging interest in material and popular culture, this trend makes perfect sense. Regrettably, libraries have never had sufficient funds to collect and preserve everything of potential research value. Thus, for libraries, this expansive view of artifactual value presents problems that are not primarily theoretical, but eminently practical.

Given the task of identifying achievable, fundable preservation strategies and goals for libraries, the task force took seriously its charge to identify parameters of artifactual value and to do so in a way that, following the spirit of preservation principles, would accept some loss as inevitable. It sought, in other words, to manage the risks of *unacceptable* loss. By looking well beyond the traditional mainstay of research library holdings—books and serials—and seeing that in the not-too-distant future the demand for analog audiovisual sources and for digital materials of all types will be as great as or greater than is the demand for print resources, the task force had to grapple with the prospect that the present preservation problem will grow to a scale that will render current approaches to preservation and access obsolete or irrelevant.

2.1 Selection for Preservation of the Original

The library preservation community has agreed on certain cardinal features of physical objects that warrant preservation in their original formats. These features are

- age
- evidential value
- aesthetic value
- scarcity
- associational value
- market value
- exhibition value

Objective criteria or established practice determine many of these features, and the criteria vary little among libraries. They are, in short, best practice. Several selection policies that are based largely on these features appear as examples in Appendix III.

The task force excluded from consideration those categories of artifacts that are always retained in the original. Such artifacts include, for example: books printed before 1801, which are usually segregated from general holdings in a rare-book collection and subject to different handling and preservation protocols; manuscripts and archival materials that exist only in few or single copies; and items with high market value. These items are often crucially important for research and teaching; at the same time, there is little debate about their value as physical objects. Moreover, their disposition after reformatting is not an issue. Despite their importance to research collections and their rightful demand on preservation resources, these types of artifacts are not discussed in detail in this report because their value is not seriously contested in the libraries that have responsibility for them.

The value of the artifact for research purposes—as opposed to its monetary value or exhibition value—is chiefly evidentiary. An artifact is of evidential value because it testifies to the extent the information in it is *original, faithful, fixed,* or *stable.*

Originality. An original manifestation of a book, photograph, or recorded performance is valuable because through it a scholar may come closer to uncovering the original intent of the creator or the publisher or both.[5] When a copy yields insufficient information about that intent, access to the original may be needed. Reformatting and copying information are analogous to translation from one language to another. Depending on the source and the target language, as well as the skill, care, and cultural biases of the translator, something inalienable to the original is always lost. A good translation, like good recopying, is one that loses the least amount of original content and intent.

Fidelity. The physical artifact is useful, and at times essential, in establishing the authenticity of an item. In other words, it has forensic value. How does one know that the item in one's hands is what it purports to be? There are internal clues in a document that give evidence of authenticity. Among these are the accuracy and appropriateness of the content. A newspaper dated 1901 that contains listings of television broadcasts, for example, is unlikely to be authentic. In addition, the external information contained in a physical manifestation provides clues of authenticity and integrity. Erasure marks on a sheet of paper, splices in a film, dried white erasure fluid on property maps—all these are physical clues to the integrity of the object and, hence, the authenticity of the information recorded in it.

Fixity. The content of the artifact when it was first produced constitutes the text (in the case of textual materials) or the document (in

[5] The instances of published versions differing from the presumed intent of the creator are legion in books, films, and other materials. In those cases, the sources that contain information about the work in prepublication form (e.g., drafts, outtakes, and proof sheets) are also required to reconstruct original intent.

the case of a photograph or an opera performance). If one is holding a fifteenth-generation fax, one cannot guarantee that the full content of the original is conveyed except by comparing it with the original, which has fixed the content by recording it at one instant in time. One of the wonders of mass reproduction of text is the way in which replication by machine en masse tended to stabilize texts that had previously been somewhat fluid. Recent humanities scholarship, however, has done much to undermine the notion of even mass printings producing a stable version or versions of various texts, and the digital realm is eroding further the concept of the fixed content of a published item.

Stability. The persistence of an object over time leads to the stable and continuous accessibility of the information contained in it. Documents whose physical substrate changes over time themselves change. Film that gets spliced and repaired loses content; digital files that get reformatted into a newer version of a software program change; photographic images printed or displayed in various manners shift tone. When one looks at a 30-year-old image of a woman in a red coat that has been printed on paper that fades and at a contemporary image made from the same negative that has been made into a slide and is being projected on a screen, chances are that the coat will appear as a different shade of red in each image. The content or value of that red is not stable, and it is difficult to efface mentally the effects of age and reformatting and to determine whether the original color of the coat was scarlet or crimson.

Some artifacts are valuable for research because the format itself is the subject of investigation. Original bindings carry evidence of print history, just as original daguerreotypes carry evidence of an early imaging technology. Bindings can also testify to the economic status of the intended reader (inexpensive as opposed to expensive presentations, for example). In these cases, the object itself is the primary source of interest.

Also of value to the research process is the physical encounter between the researcher and the object—an encounter that can help prime the scholar's imaginative and analytical skills. While this factor is highly subjective and difficult to quantify, many scholars claim it has had, at least at some stage of their careers, an irreplaceable heuristic value. A medievalist who has never worked directly from manuscripts is at a disadvantage, just as a biographer of Thomas Jefferson who has worked exclusively from the printed editions of his letters may be said to work at one critical remove from his or her subject. Nevertheless, given the toll that physical handling takes on all types of materials, the task force considered that surrogates can be judiciously used by those who have a familiarity with original source materials and that, from the perspectives of preservation and convenience of access, surrogates are often preferable.[6]

[6] See, for example, the case of the international editorial team working on the James Boswell Papers at the Beinecke Library (Bouché 1999). The editors came to prefer using digital scans of the original manuscripts to working from the originals in New Haven. In part this was a matter of convenience—the work could be done wherever the editors were located and obviated the need for repeated trips to New Haven; however, the editors also recognized that the scans were more legible than were the originals.

2.2 Frameworks for Determining Value

Questions about the nature of the artifact have caused scholars and library professionals to realize that, even for the early part of the nineteenth century, much more information of potential research value exists in traditional formats such as paper and image than had previously been recognized. Consequently, the process of redefining what constitutes an artifact must be done not only for new media but also for a considerable body of information from the 1800s. The fragility of paper-based materials printed since 1850, especially newspapers, has been a concern for some time (Baker 2000, Cox 2000, Marley 1975, Smith 1995). Because newsprint is so fragile, preserving one or more instances of all imprints of newspapers poses enormous technical as well as financial challenges. More recently, there has been a growing awareness of other kinds of artifacts from the nineteenth and early twentieth centuries that also require the attention of preservationists. These include materials that serve as primary evidence of popular culture, ranging from baseball cards to ladies' magazines to dime novels to political posters. All were produced for a mass market using cheap and readily available materials.

The examples from popular culture raised the most difficult matter the task force considered: How to weigh the relative merits of various claims on scarce preservation resources. The need to consider merit is nothing new. Scholars, archivists, and librarians have always assumed a hierarchy in collections. The artifact or original document was the item initially collected and accessioned. When it was prized for some unique aspect of its material or historical existence (e.g., a first edition, a holograph manuscript, a signed author's copy, or a presentation copy of a work), the artifact was given priority for preservation. Absent that uniqueness, a lower level of care might be given, and a lost or damaged copy might be replaced by one whose physical manifestations were quite different (e.g., a photocopy or a reprint). The value of the unique artifact could be defined variously as "historically important," "rare," "associational," and so on. In each instance, however, there was an individual material object that someone had once defined as valuable enough to retain.

The Achilles' heel of traditional definitions of the artifact lay in the value judgment that determined artifactual status in the first instance. What were the grounds for deciding in favor of one object and against another? How can libraries cope with the fact that the value of the artifact is never quite the same to different researchers? While one scholar will seek certain information from the item, scholars from other disciplines will require different kinds of data that may involve a wholly different way of handling the object.[7] Can one say that these users are even seeing the same object?[8]

[7] Research disciplines train scholars to attend to the materiality of their objects of study in very different ways. Historians, literary scholars, philosophers, art historians, historians of science, linguists, and text editors examine quite different aspects of an object of study, which, in the case of artifactual objects, will be complex. Sometimes there may be overlap, but more often there is not. The needs of all serious users are legitimate, and preservation should serve them as effectively and equitably as possible.

BEST COPY AVAILABLE

Scholars and others who use artifacts may think of them in the aggregate as unified objects; however, when they interrogate the artifact for their research, they tend to focus on parts rather than the whole. One may define the artifact as a series of multiple discrete components—handwriting, watermarks, marginalia, splices, evidence of use—each potentially a focal point for scholars and others, depending on what they are studying. From the standpoint of usage, people normally analyze discrete sets of information contained within an artifact. The fact that artifacts are complex and that they lend themselves to a variety of intellectual endeavors means that one must think of them in terms of their parts, and not just as wholes. One way to think of the artifact, then, would be as a multiplicity of informational sets, including the material form of the object and its contextual history, where known. (The relationship of part to whole is also crucially important.)

One element of the artifact's complexity is the fact that scholars from different fields will perceive and use it in different ways. A first edition of a novel by Charles Dickens will be used very differently by a historian of Victorian England researching the economics of the book trade, a literary scholar concerned with different versions of the work, an art historian interested in Victorian book illustration, a textual historian interested in layout, and a historian of bookmaking. Each will consult the same artifact for a different kind of information, and none may notice the information sought by the others. Similarly, photographs of the Civil War by Alexander Gardner can be used to study the battles; the public's reception of the war in the North; the history of clothing, medicine, or gender; or even the medium of photography. Which subject interest and methodology would require use of the original? Which could make do with copy prints? Which could make use only of the original photographs in their original presentation portfolios? Which would be enhanced by access to the images through digital delivery, which could then be manipulated to magnify details?

The preceding observations suggest the possibility of proposing a contextual definition of the artifact as follows:

> An artifact is a physical object produced at some time in the past, and attesting to a given set of practices, thinking, and ways of viewing the world, and whose importance will be defined by present and future needs and use. The value of the artifact is strongly influenced, but not completely determined, by its rare or unique features.

The artifact conveys historical consciousness in different ways, depending on who studies it and for what purpose. Much of the information conveyed by an artifact does not require the presence of

[8] Artifacts derive their value from how they are viewed and used in a given culture at a particular moment. As cultural variables, they will be viewed and studied differently in different periods. A holograph copy of a speech by Robespierre would have a different value for a royalist in Louis XVIII's government in 1816 than for a socialist historian in France in the 1990s.

the physical object. Surrogates of the object (e.g., photographs, photocopies, and digital versions) may convey much of the information stored by the artifact. Indeed, for many purposes, a high-quality surrogate may convey this information better than does the original. The surrogate may enable access and use that would otherwise be impossible; for example, it allows a user to view an object that is physically distant, to enhance images, or to perform full-text searches. Surrogates do not obviate some scholars' need to consult the object itself; however, in many instances, a surrogate can serve scholarly needs as well as, or better than, does the artifact itself.

The artifact matters. It matters very much. Nevertheless, in a time when artifacts are abundant and resources scarce, the scholarly and library communities are called to rethink the status of the artifact in terms of its content and material form. Acknowledging that every aspect of an artifact yields information that will be of use to some scholar or other, we nonetheless need to assess the relative importance of the different aspects of an artifact pragmatically. Scholars and library professionals jointly face issues such as the following.

First, within the timeframe of the last 200 years, what constitutes an artifact worth retaining? The answer is not obvious in the case of nineteenth-century material artifacts: baseball manuals or railway timetables have not traditionally been viewed as important cultural documents, although they probably would be today. The question becomes truly perplexing in the case of media that are dependent on playback equipment, from recorded sound to moving images, in which the concepts of original or unique, stable or fixed, may not even apply.

A second question concerns resources and priorities. This imperative may be spelled out as the "how, who, and when" of artifact preservation. In other words, accepting the reality that resources for preservation are limited, and assuming some common, gross-level understanding of the value of artifacts, how (in what form) are they to be preserved, by whom, and when (or how often)? Are all libraries or archives to be held responsible for collecting and preserving the same categories of artifacts? How much redundancy of preservation is necessary? How much can libraries afford? Redundant collections serve as insurance policies for preserving and making accessible information in a physical format. This is even true for digital information, which can be cloned and shipped around the globe almost instantaneously, yet is highly fragile if no one assumes archiving responsibility for the data. In the case of digital information, given the high costs of building and maintaining information technological infrastructures, the same degrees of redundancy that we see for print may be neither desirable nor feasible. If and when digital files become the default mode for access—even for materials such as journal articles or encyclopedias, which were originally physical artifacts—what are the implications for duplicative collections of physical artifacts?

The point is not to determine whether baseball manuals and railway timetables from the nineteenth century, for example, are of suffi-

cient artifactual value to justify the expense of collecting and preserving them. Rather, it is to set priorities in the face of financial constraints that too often mean, practically speaking, that if one decides to collect and preserve one kind of artifact, resources for other kinds will be insufficient.

The issue is not to evaluate the artifact per se to determine what survives and what does not. The scholarly community has no more of a claim to the wisdom of the ages than does the library community. The issue is to identify productive methods for interrogating the individual artifact that would, in a climate of finite resources, inform decisions about whether and how to preserve it. Such methods would help ensure survival of the greatest number of artifacts by intelligent analysis and careful consideration among knowledgeable and committed communities.

There is no single method of engaging primary sources, nor is there one overriding set of priorities—text over image, manuscript over map, English language over non-English—that will meet the needs of present or future research. But how do scholars, librarians, and archivists work together to prevent the kinds of losses that we can now regret in leisure—local newspapers, silent films, early television broadcasts, among others? Some past initiatives have brought together the scholarly, archival, and libraries communities to propose preservation and collection guidelines. The Commission on Preservation and Access convened scholars in the 1980s to identify brittle books to microfilm (George 1995). In the 1990s, the American Historical Association, the ARL, and the MLA created a task force on the preservation of the artifact to document preservation challenges and inform scholars about the need to collaborate with libraries to address them (Reed-Scott 1999). These efforts have effectively addressed specific problems. However, most of the collaborations between scholars and librarians have been either locally based or designed to address only specific areas of concern, and they have focused largely on print materials. One goal of the task force was to propose strategies for collaboration that are realistic and sustainable, that balance the needs of present and future users, and that address the proliferation of nonprint materials in the information landscape.

American scholarship has historically relied on a decentralized network of libraries, archives, historical societies, and museums to collect, preserve, and make accessible the intellectual and cultural resources that form the basis of academic inquiry. This decentralized and largely uncoordinated approach is unlikely to change, despite the resulting losses of artifacts that have occurred. However, new preservation and information technologies promise to offer cost-effective means to prevent or slow physical deterioration and to keep the research community well informed about what others in the community are doing to preserve their holdings. These technologies will be successful to the extent that the research community is committed to identifying preservation challenges and to lobbying for financial resources—from provosts and foundations to federal agencies—to address them. Strategies for collaboration will depend on

the use of new information technologies to share crucial information about the status of various collections. They will also depend on the continued engagement of scholars, librarians, and archivists to deploy those technologies in meeting their responsibilities to the well-being of the intellectual and cultural assets whose benefits they enjoy.

3. States of the Artifact 1800–2000

3.1 The Changing Features of the Artifact

Research library collections are primarily made up of printed matter: books, serials, journals, and newspapers. Even with the rapid growth of machine-readable and electronic resources, 85 to 90 percent of the acquisitions budgets of libraries still goes toward the purchase of printed matter.[9] Humanities and social science scholars rely heavily on print journals for scholarly communication and so do scientists. Those who work within the print regime rely on a series of conventions about documents and their relationship to their physical manifestation that may be so familiar that they are invisible. These conventions bear mentioning, however, because many of them do not operate in realms of audiovisual and digital resources.

Among the many consequences of the adoption of printing technology, several became fixtures in the print landscape. They are

- the creation of a comparatively fixed and stable text
- the concept of authorship and of intellectual property inhering in authorship
- easy duplication and wide dissemination of texts, especially after the introduction of high-volume presses in the nineteenth century
- the notion of fungibility of informational content (also a result of mass-market publishing)

While texts copied by hand were intended to replicate their sources, they did not do so precisely or completely. With the advent of printing, it became possible to produce nearly identical copies in large numbers. This great increase in the accuracy of reproduction was crucial in the development of scientific and technical literature,

[9] In 1998–1999, both ARL and Digital Library Federation member libraries spent about 10 percent of their resources budgets on electronic resources (Jewell 2001, 4).

and especially so in the reproduction of illustrations. While there certainly have been variations between printings, and even among copies from the same press run, the presumption that authorship and content were, in principle, stable and fixed, took hold, with significant consequences. The notion of repeatability and accurate reproduction, on the one hand, and of a reliable text with a known author, on the other, caused a shift that Foucault describes in detail in "What Is an Author?" According to Foucault, with the arrival of printing, scientific literature ceased to derive its claim to legitimacy from its attribution to an author (Aristotle, for example) and began to derive its authority from principles of experimentation—the potential falsifiability of the hypothesis, the repeatability of the experiment, the replicability of the result. Literature, on the other hand, began to derive its authority not from its dispersion and repetition in the culture, but from its originality and its connection with a particular author (Foucault 1977). In an information economy, much now depends on these notions of fixity, reproducibility, and authorship, as many recent court cases, legislative acts, and international agreements attest.

Bibliographic and textual scholarship, since the nineteenth century, has shown just how precarious and nuanced these concepts are. The destabilization of our ideas of fixed content began to accelerate, however, with the advent of audiovisual recording technologies. Daguerreotypes exist in only one version, because the image is exposed directly onto a metallic plate; wax cylinder recordings are also unique, each produced as a live recording. (Although they were produced in batches, there was no master.) Film-based images, however, do have a master—the photographic negative. For this reason, they can exist in multiple, nearly identical copies. Still, although the images are all made from one source (i.e., a negative), the negative inevitably wears with use, and the copies become less faithful. Through reproduction, the image on the negative can be effaced. This is not as true of print products. Engravings do deteriorate with each production, and books set with type can show wear from copy to copy; nevertheless, the loss of information for texts from copy to copy is generally not as striking as it is for visual and sound resources.

In non-print materials, the problem of "version control"—i.e., determining which version should be preserved as an original—is complicated by a further lack of stability over time and place. This factor characterizes broadcast media in particular. Take as an example a single television document that a news archives might wish to collect: the CNN news broadcast for January 15, 1998. What do the archives collect? The broadcast that originated in Atlanta, in London, or in Tokyo? This problem has existed among collectors of newspapers, of course, with libraries deciding often on the final edition as being the one of record and making efforts to collect several editions of a paper that tracks a particularly important event, such as an assassination or an election. But the scale of this problem for broadcast media on a 24-hour news cycle that constantly update news is of a different order.

The artifact that records an image or sound, moreover, can easily lose its originality or uniqueness. The adjective "unique" may have once been sufficient to identify a primary document in manuscript, and in some cases, such as legal documents or signed first editions, it may still have some meaning. But with newer formats, such as electronic and broadcast media that rely on refreshing and reformatting for longevity, the terms "original," "unique," "content fixity," and "material artifact" mean little.

Task force members found the cardinal features of an artifact that have the highest research value—originality, faithfulness, fixity, and stability—retreating like a mirage as they worked their way from the 1800s into the twenty-first century. The three sections that follow look at how these four values work themselves out in print, analog audiovisual, and digital documents. These sections also explore how preservation strategies address these values and seek to preserve them. Preservation options, it will be seen, are shaped by such factors as the quantity and quality of resources, the instability of media, constraints of resource allocation, and the changing valuation of the artifact in research and teaching.

3.2 Print/Paper

3.2.1 The Relative Stability of Imprints

Books and other printed matter deteriorate over time as the result of their inherent chemical instability. For example, when paper made of wood pulp reacts with humidity and heat, it becomes brittle. Books also deteriorate as a result of mechanical strain. For example, the spine is stressed when an open book is placed on a photocopy machine. (Photocopying also exposes the paper to heat and light.) But there have been significant changes in how books, journals, and newspapers have been made since 1800, and these changes affect their significance as artifacts as well as their physical robustness.

In the 1970s, 1980s, and 1990s, much attention was focused on the legacy of brittle books created by the processes for making inexpensive paper. Now joined to this concern about the paper is concern about the structural support for that text block—the binding. The last few decades have seen an explosion of relatively inexpensive, softcover editions of books that were not designed to last. Thus, the task force looked at both chemical and mechanical problems associated with paper and its binding.

Paper. Until the middle of the nineteenth century, paper was made from linen and cotton rags, which in principle make a strong and durable product.[10] In the 1850s, wood pulp came into general usage for making paper more economically. The publishing industry rapidly converted to this process, following the lead of the newspapers, for which wood pulp was a source of inexpensive newsprint.

[10] Acid was sometimes used even in the production of rag paper. The first problems with acid paper appeared in the 1830s, when rag paper was bleached with acid-producing chemicals that weakened it.

The manufacturing process required that wood-pulp paper be treated with aluminum sulfate (alum) to keep the inks from running and to improve the hand. Alum, together with various bleaches and sizings usually added during the papermaking process, reacts with humidity to produce an acid that, over time, breaks down the molecular structure of the cellulose in the wood pulp. In its worst form, damage leads to "brittle" paper that loses its flexibility and eventually crumbles when handled. While any given page may be readable, turning it may lead to several types of damage. In thin, hard-finish paper, the page becomes brittle and brown along the edges and can easily snap off along a fold line. Thick, pulpy paper tends to separate almost spontaneously under tensile stress in any direction, and whole blocks of text may come loose near the gutter and fall out of the volume (Kantor 1986).

Once paper becomes embrittled, nothing can arrest the decay. Such materials are candidates for preservation reformatting, that is, capturing the information content of the original and transferring it onto a stable medium such as preservation-quality microfilm.[11]

While acid paper of any sort is at risk, decay manifests itself unevenly. Manufacturing processes vary a great deal, and the conditions of storage and use can vary dramatically as well. Few places have proved to be as bad for wood-pulp paper as the humid, polluted eastern seaboard of the United States. Books, like people, prefer California-like climates that are temperate and do not vary dramatically.

Many acid-paper items are not yet embrittled and can be stabilized to arrest the process of embrittlement. The two chief methods of stabilizing acid-based paper are deacidification and storage under optimal conditions. Deacidification is a chemical process whereby paper is treated with an alkaline buffering agent that neutralizes the acid content. It can be done on a single item or on many items at a time. There are facilities that can deacidify bound materials en masse, and some facilities can even treat unbound materials, such as sheet music, archival documents, and newspapers. Deacidification can stop further damage, but it neither reverses damage done nor strengthens already-brittle paper. Therefore, it is unsuitable for books that are damaged or weak. Part of the cost of treatment lies in carefully assessing, item by item, how suitable a book is for deacidification.[12]

Embrittlement can also be avoided or slowed by storing materials in stable environments with set parameters for temperature and humidity. These storage conditions cannot be obtained in open-access stacks. The conditions that slow the decay of library materials

[11] Not all microfilm is created equal, and the preservation profession recognizes certain standards of film capture, film quality, and storage as being preservation-worthy. Unfortunately, a significant number of microfilms do not meet these standards, resulting in surrogates of poor quality.

[12] A relatively new technology, known as paper splitting, strengthens weakened paper by splitting a piece of paper in half to separate the front from the back, inserting a stabilizing layer, and reapplying the two halves of the paper. This process, which was developed in Germany, is not widely deployed in the United States.

would be uninhabitable for most people. One of the reasons that libraries have been eager to build remote storage facilities is to lengthen the productive lives of their print collections.

How extensive is the problem of brittle books? In 1984, the Library of Congress and Yale University surveyed their holdings and found that one-quarter to one-third of their collections were highly embrittled and in danger of imminent disintegration. This alarmed other libraries, which turned to their partners in the academic community to help assess their collections and devise a coordinated strategy to address the problem of brittle books. The Council on Library Resources asked the Association of American Universities and the American Council of Learned Societies to join in creating a task force to study the extent of book deterioration. In 1984, Robert Hayes, then dean of the Graduate School of Library and Information Sciences at the University of California at Los Angeles, was commissioned to determine the percentage of embrittlement at major repositories in the United States. He determined that of the 305 million volumes in Association of Research Libraries collections, approximately 25 percent were brittle. Hayes also attempted to determine the degree of overlap among libraries to find out how many of these were individual titles that needed to be preserved. The number he arrived at was 12 million, and he estimated that about one-third of these could be microfilmed in a 20-year period (Hayes 1985).

The reformatting of brittle books accomplished two things: it rescued information deemed endangered and increased access to that information—a point that was critical in persuading the U.S. Congress to fund the National Endowment for the Humanities (NEH) Brittle Books Program. Each reel of preservation microfilm produced under the auspices of NEH was made available for purchase, in accordance with any copyright considerations, and each title filmed was entered into a database that recorded the existence of the film. This strategy not only helped avoid accidental duplication of effort but also made known the availability of the titles.

Although important for meeting the needs of remote users, microfilming books seldom best serves the access needs of local users. Photocopying onto acid-free paper is the preferred technique for this purpose, and it is an option used increasingly by most libraries.[13]

But what does shared access to the artifact look like? How can a single book serve the needs of both local and remote users? The example of registering reformatted books raises a question for those engaged in mass-deacidification programs. Does each library have to duplicate the deacidification work of the others? Would it be possible to create national registries where libraries log the books that they have treated? Other libraries could then consult the log and determine whether local demand dictated treating their copy of the work, or whether it would be acceptable to box or shrink-wrap the work and send it to remote storage, knowing that if that copy became too

[13] Johns Hopkins University is doing more preservation photocopying and less microfilming, and this trend is common among ARL libraries. (Testimony of James Neal to the task force, October 29, 1999.)

BEST COPY AVAILABLE

brittle to use, library patrons would have access to another copy through interlibrary loan. Unlike reformatting on film or digital files, saving one book in its original form does not increase access to it. The problem is, how can libraries achieve economies of scale in the preservation of artifacts? The first step would be to improve passive systems, especially environmental conditions. If libraries do not intervene to save every low-use book that will turn brittle, but take action (such as deacidification) to stabilize some number of them that can fulfill the needs of patrons who must use an original, how many of a single item should be preserved, where, and at whose expense?

Yale University's Sterling Memorial Library has microfilmed more than 60,000 books during the last 12 years. During that time, though, the library acquired 150,000 additional books each year, and more than 65 percent of these came from countries where permanent, or alkaline-buffered, paper is not used. In other words, almost 100,000 volumes being added to the library's collection each year are at risk of becoming brittle in the future. On the basis of its own estimates, Yale determined that filming a volume after it became brittle would cost about $120 in current dollars. Scanning the volume might cost as little as $80, but the cost of preserving and managing digital files over time is unknown, and, in any event, it is not yet a preservation format. Deacidification would cost $17 a volume. This is one library's estimate of how cheap an ounce of prevention would be in comparison with a pound of cure (Walls 2000). This is in line with the LC's estimates of the various treatments available for print materials (see Appendix VI).

With respect to the future, the good news is that publishers in the industrialized world had largely ceased to use acid paper by the 1990s, at least for first printings. In 1990, the U.S. government mandated the use of permanent, acid-free paper in all documents and publications that were to be archived, and most state governments have followed suit. Most publishers agreed to print the first press runs of hardcover books on permanent paper.

Librarians have noted a disturbing trend during routine checking of new acquisitions for acidity. In many newly independent and emerging countries—countries from which U.S. libraries get a significant portion of their collections—printers continue to use unbuffered wood-pulp paper. A large portion of these acquisitions are at high risk for acidity. But more and more academic and trade presses in the developed countries are producing reprints—and an ever-increasing number of first press runs in paperback—on acid paper. Consequently, many new acquisitions, not only from Asia or South America but also from Europe and North America, may be at risk of embrittlement unless deacidification or other preventive measures are undertaken.

Binding. In the West, changes in the publishing economy have created yet more preservation problems. To save money and cater to mass markets, publishers have sharply increased the numbers of paperback books published. As a result of this trend, which began in the 1950s, libraries need to rebind these acquisitions even before put-

ting them on the shelves. This may divert money that would otherwise be spent for preserving older items.

Books and other bound materials are vulnerable not only because the paper may be weak but also because physical handling weakens the structure of the volume and creates other problems. A study done in 1994 at New York University yielded some interesting data. During one week, preservation experts looked at everything that came across the circulation desk. They found that 21 percent of circulating books returned with spine damage and 14 percent needed rebinding or recasing. Thirty-two percent of the books returned were stained or damaged in some way.[14]

The latest ARL preservation statistics (1996–1997) reflect the local need to keep books as objects alive and well. ARL libraries repaired 12 times as many books as they filmed (873,000 versus 70,597) during that period. Thirty-six percent of total preservation expenses went for contract binding; filming accounted for only 3.6 percent. Of the 2 to 4 percent of library budgets that go to preservation (exclusive of capital costs), item-level repair appears to be the asset-management strategy of choice.

3.2.2 Evaluation of the Artifact and Selection for Preservation

The recent interest in the production of knowledge has made materials that libraries acquired as secondary sources in the last two centuries into primary sources for this century. The lively field of printing history has made the book itself, rather than its content, the subject of investigation. This has promoted a large and not very physically robust category of resources to artifact status.

It is not surprising that historians of print are interested in books that were created before the mechanization of printing and binding. Book production before the 1830s was craft work. Printers delivered unbound sheets to the booksellers, who then bound them into volumes that carried the printers' own imprints. One printing was distributed among a number of booksellers, each of which would bind the volumes differently. Consequently, the same printing could appear as a number of forms. Each set can be considered unique in its printing, binding, or dissemination and thus worthy of retention in a collection.

The introduction of mass production has not diminished the status of books from the latter two-thirds of the nineteenth century as items with artifactual interest. There were so many innovations in the business and art of printing during that period that there are now many candidates for special treatment as objects, irrespective of the content. While books printed before 1801 are usually managed as "rare" books, there is an increasing use of the category of "medium-rare" books of the nineteenth century that are served to patrons un-

[14] Testimony of Carol Mandel to the task force, October 29, 1999. A similar study of circulating collections at several Columbia University libraries in 1995 revealed significantly higher levels of embrittlement and other damage in the oldest collections. More funds were subsequently allocated for rebinding and protective enclosures. (Personal communication, Janet Gertz to Abby Smith, April 4, 2001.)

der somewhat stricter protocols than are general collection imprints. This could be because the books have illustrations that are of research value or are vulnerable to mutilation; have aesthetically significant bindings; or were produced by special printers or publishing houses. These books rarely receive the intensive, item-level treatment that a rare book would get; a global or collection-level treatment is sufficient to ensure that they survive in usable form. This is one example of preventive preservation that should be encouraged in libraries holding important collections of such materials.

One should not underestimate the amount of time it takes to select books for this kind of treatment. (Books are rarely shelved by age, after all.) Some libraries pull these books during normal shelf inventorying or when they cross the circulation desk. Other libraries pull these books when they are selecting items for filming, scanning, deacidification, or secondary storage. To expect each library to develop a collection of artifacts relevant to the history of printing simply because it has books from eras relevant to that field, however, is neither feasible nor responsible.

Many paper-based collections other than books and periodical publications held in special collections libraries and departments are also at risk. Libraries, archives, historical societies, and museums house large collections of pamphlets, letters, brochures, broadsides, sheet music, printed maps, advertising art, playbills, restaurant menus, scrapbooks and memorabilia, almanacs, proof sheets, children's books, religious tracts, and other items printed between 1800 and 2000. The task force has not focused on these collections because, to the extent that they are rare or unique and constitute primary sources, institutions that hold them do not dispose of them. Nevertheless, retention of these materials does not guarantee their survival. These collections are invaluable intellectual and cultural resources, and they must be considered in any national preservation strategy. While these materials fall outside the scope of consideration here, it is important to emphasize that these collections warrant separate investigation. Indeed, the ARL recently completed a study of special collections within its membership that underscores the need to devote attention to the stewardship of these types of collections (Panitch 2001). The report points to a number of areas, including but not limited to preservation, that demand fresh approaches as well as new resources. The task force recommends that such studies be extended to non-ARL libraries that hold special collections, that the needs of these collections be identified, and that a strategy for devising and funding cost-effective solutions be developed.

3.2.3 Creating Surrogates: Filming versus Scanning

Surrogates are created for one of two reasons: to create copies of works too fragile to use or to replace items in imminent danger of disintegration. In the former case, a rare book or collection of broadsides may be scanned and the originals retired from use, except under extraordinary circumstances. For materials that are on their last or next-to-last use, the content is reformatted onto a more stable me-

dium, such as preservation microfilm, or is photocopied onto acid-free paper. For these types of books and other items that are in current demand, most libraries create paper copies, a far more convenient mode of access than is microfilm. In most cases, the source is photocopied onto acid-free paper; however, at libraries with active digitization programs, the source materials are often scanned and the scans are used to recreate the original volume on acid-free paper on demand. The question that then arises is what to do with the original. This issue is addressed in Section 3.2.4.

Microfilming remains the gold standard for preservation reformatting of low-use materials. With proper storage and handling, preservation microfilm can remain faithful and legible for a century or more. Film is still considered to be the best medium for preservation of images. But microfilm is just that—images—and no more. Digital reformatting that includes optical character recognition (OCR) adds functionality, including the capability for full-text searching.

Given the value added by digitization, why isn't all reformatting digital? The chief reason, besides the often higher cost of scanning and creating searchable text compared with filming, is that there is as yet no reason to be confident that digital files will last as long as microfilms, or be as easy to manage over time.[15] The preservation community has given much thought to making preservation microfilm whenever digital scans are made or to converting preservation microfilm into digital scans (Chapman, Conway, and Kenney 1999), but few libraries have embraced this more costly approach. Even though the money goes toward ensuring preservation after creating access, funders tend to be reluctant to put money into preservation when the same money can be put toward enhancing access to something else.

The investment taxpayers have made, through the NEH, to create microfilm copies of brittle books could be leveraged to create digital scans of those books for ready access. This notion, however, has elicited little enthusiasm. Part of the reason is that the books that have been microfilmed are often low-use items. As one expert has written, "Brittle books have been selected for filming because they have potential research value, but are low priority for current researchers and so can be put on film for storage even though it is an awkward access medium" (Gertz 1999). Digitization of low-use special-format collections, by contrast, is common. This predilection to scan special collections rather than monographs is based in part on the idea that special-collections materials—maps, photographs, manuscripts—have traditionally received little use because they exist in single copies in one collection. Once made easily accessible, these materials may become high-use items.

[15] For a full discussion of the technical, legal, financial, and administrative complexities of ensuring the longevity and integrity of digital files, see Task Force on Archiving of Digital Information 1996. The problem of preserving digital information is treated briefly in Section 3.4 of this report.

One significant exception to the practice of not scanning low-use print materials is the Making of America (MOA) program at the University of Michigan and Cornell University. This digitization program focused on brittle American monographs and journals from the latter half of the nineteenth century, and the selection of materials was coordinated between the two libraries. This conversion of brittle American monographs and periodicals has created an interesting and largely unanticipated result: the MOA journals at Cornell receive thousands of hits weekly. The beneficiaries of the project are many. For example, a graduate student used the publications as research materials for his dissertation. They enabled him to complete his studies from abroad when his wife's job took them out of the United States. William Safire noted the value of mining the text for early uses of words and phrases and cited MOA publications scanned using OCR as a rich source for nineteenth-century texts (Safire 2000). A high school student found the ideal material for a term paper, while a company in Detroit located an engraving of Daniel Boone, which they intend to make into a poster to commemorate the city's 300th anniversary. The visibility of digitized materials on the Web has facilitated their discovery, resulting in usage that greatly outstrips that of the paper copies that were slowly decaying on their shelves. A dedicated researcher may have consulted the print copies, but the secondary school student, the genealogist, the lexicographer, and the insurance company are unlikely to have engaged with the paper volumes as they are able to access these digitized publications on the Internet.[16]

A recent study on how humanities scholars work in this evolving information environment, based on surveys and case studies, reports a similar trend in preferences for electronic access to print materials via Web-based data repositories:

> . . . the scholars all had access to a number of full-text databases published by Chadwyck-Healey, the Women Writers Online project from Brown University, and full-text journals from the Johns Hopkins Muse project and from JSTOR. Few scholars mentioned using these full-text resources, but the ones who had were hooked on what they could offer and particularly appreciated those products that provided access to primary sources (Brockman et al. 2001, forthcoming).

The scholars who were "hooked" reported that the searching techniques available were especially prized. The report noted that "the thoroughness with which searching is possible across any of the corpora covered by these databases means that once they have been recognized by a group of researchers in a particular field, their use is obligatory."

The observation that use is "obligatory" means that these scholars are now able to avail themselves of otherwise-scarce texts. Wom-

[16] This transformative effect is further borne out by JSTOR, another project that is building a dynamic database from the static pages of journals. For more information, see Section 4.4.

en Writers Online was cited specifically for both research and teaching uses. The types of searching are novel as well, being inaccessible to manual research; they include such techniques as keyword searching, pattern identification, and an abundance of searchable elements. Finally, the database created a contextual mass of different editions of the same work that allows collation and comparison among versions. In addition to reporting use of the larger commercial databases, a few scholars recounted using smaller, noncommercial Web-based projects devoted to individual authors.

All these instances testify that scholars are growing more comfortable with digital surrogates of texts. There is a need to develop and apply methodologies to track the growing use of large digitized collections and to evaluate how researchers use these aggregations. Their increasing use also raises the question of what value there will be for libraries to maintain their hard-copy collections of these texts, when the same collections, without gaps and items missing from the shelves, are accessible from a computer 24 hours a day from anywhere in the world. Duplicative collections of materials that are not rare and are readily accessible on the desktop will have lost much of their original reason for being: to provide local access. What will the next generation of scholars, many of whom will have grown up with JSTOR, MOA, and other databases and find the very notion of having to retrieve hard-copy journals from stacks during library hours a hindrance to research, think of the journals and monographs that languish in dead storage? How much will a university or college be willing to pay for digital access and also for keeping hard copies? How do we ensure that enough originals are kept and are available for present and future use?

Although we are far from having a comprehensive body of primary and secondary literature readily available for use from the desktop, there is every reason to believe that within a decade, significant corpora of texts will be available in a number of fields. It is not too early to plan for the eventual disposition of the scores, or even hundreds, of duplicate copies of individual items that scholars, voting with mouse clicks, prefer to use online. The question is not whether libraries should keep hard copies of them. Of course they should. The real question is, in a world in which access is no longer tied to physical possession, let alone ownership, which institutions, and how many institutions, will take the responsibility of ensuring the preservation of and access to hard copies? Information technologies may allow libraries to develop a system of registration or other kind of tracking that can allow linking local preservation decisions and investments to the changing needs of a national and international research community (Greenstein 2001). And with new models of distributed preservation responsibilities facilitated by digital technology, shared among libraries, must come financial and legal support for those actions taken on behalf of the group.

3.2.4 Responsible Retention Policies

Books are totemic objects that have developed a powerful status over centuries, both for their content and for their physical significance. At the same time, books, like other objects, have always been cultural variables. Codices were routinely expunged in the Middle Ages, their writing scraped off to make room for a new work, which, in turn, might be effaced by a later generation. That said, task force members testified repeatedly to the symbolic value of the book and to the cultural significance of the library both as a building and as a mental construct. They asserted that responsible stewardship is necessary to strengthen and reinforce those values. The spiraling costs of basic library functions and the added demands of new and more expensive services and collections cannot be allowed to put libraries and their missions at risk. To ensure the continued accessibility of current library collections into the future, not to mention the extension of access to those collections through networked resources, economies of scale, some of which entail the forging of cooperative enterprises, must be achieved. The technical infrastructure, financial resources, and public support must also be secured to sustain those actions.

As a way of envisaging how responsible retention of reformatted materials might be reconciled with the economic realities of preservation, librarians can look to archivists for sound guidance (Kenney 1996). The benchmarks of sufficient quality film or digital capture include the following:

- The scan has captured all informational content—color, original formatting, full content, or whatever else is important in the source document.
- The document is fully accessible through a defined indexing-and-retrieval scheme.
- Access to the digital file will be maintained over time, and the data will be protected from loss, corruption, and technological obsolescence.

A consortium of leading American research libraries has proposed benchmarks for digital capture of text and image for libraries that posit technical specifications that would meet these requirements. If adopted and widely practiced, these would go a long way to building digital files on which researchers could rely for some minimal levels of capture and fidelity (DLF 2001a).

Once a library has created scans of sufficient quality to serve as full surrogates, it should put in place a plan to maintain those resources over time. It should also consider the question of what to do with the original source material. A plan for maintaining the files over time may one day become routine, but at present there are very few libraries or archives that can or would assert that their digital assets are secure for more than a few years hence. There is a need for what might be called digital service bureaus or utilities that would provide such services as scanning, managing files, delivering files to local users, and long-term archiving. This infrastructure needs to be

in place before most libraries and archives could develop routine, cost-effective digital services.

Any scanned material that is rare and of artifactual value should be handled with care during scanning, and retained afterward, even if retired from active duty, and stored to prevent further damage. Items that are common, such as journals, and that have content value but little artifactual value may also be sent to storage. However, if hard copies exist at other sites, there is no compelling reason to retain them, unless the local patrons have a history of using hard copy even when digital files are available. What is important is that researchers who do need hard copy should be able to locate and retrieve it with relative convenience. In many cases, this may mean developing shared repositories for originals, used and supported by several libraries that store their materials in one site and are able, thereby, to create richer and often more comprehensive shared collections. The custodians of artifacts need to design a plan that strives for the most comprehensive coverage of given journals or subject matter. This would almost certainly entail collaboration by the repositories of the original materials, support from their local constituencies, and concurrence that acting for the greater good might engender some local inconvenience. The Five-College Library Depository (see Section 4.1) is a model for this kind of shared storage. Such a repository might be able to afford collections services such as in-house preservation facilities, or even scan-on-demand services, that single institutions could not develop and sustain alone.

Promoting better understanding of the importance of the artifact will require a clear, succinct framing of the issues and structured discussions. The objective should be to develop commonly understood and widely endorsed approaches to the problem of caring for an abundance of materials with limited resources. There will never be complete agreement: the matter of what to preserve and how to make that accessible will never be considered resolved for all time. Instead, interested parties should develop and declare a basis for making those decisions that must be made in the immediate future.

The task force's investigation of the print record has confirmed that there is a great need for the identification and preservation of numbers of important artifacts. Materials that are unique can be identified and preserved. But there are large categories of printed materials that exist in abundance and do not have high market or exhibition value, but are important nonetheless because they exemplify a genre and a way of recording information and communicating. For these materials, the need for the identification is especially great. Such things as antislavery pamphlets, election broadsides, and sermons printed as circulars are distinct genres that have value but are unlikely to survive in toto. A measured plan to save numbers of such genres will probably succeed in securing funding for preservation; a general alarm about their status probably will not. It is important that scholars work with librarians to identify and define categories of materials and locate the finest and best-preserved specimens.

Beyond specific genres of printed ephemera, there is a need for a repository of American imprints—what one of the task force advisers called a "Federal Reserve of National Literature"—that would ensure that an archival copy of American imprints is preserved. The American Antiquarian Society (AAS) has been working for decades to create a deposit of record for American imprints up to 1876. Other than AAS, there is no library or consortium of libraries attempting to preserve a record of American imprints. The Library of Congress, despite the fact that it receives two copies of every copyrighted book in America, does not have a program, or indeed a mandate, to preserve even one copy for posterity. It is worth exploring the possibility of seeking congressional authority and funding to require the Library of Congress to send one of its two copies directly into storage upon receipt and to be obliged to make a surrogate of that copy available for use if the title becomes too scarce to find elsewhere. The Library of Congress has congressional authorization to build additional storage for its needs, but does not have the authority to set aside a copy for archival retention.

The Library of Congress is the working library of the legislative branch of the United States government, and its collections are at the service of the Congress. A request for an archival deposit system has never been made to Congress, and members of Congress neither know about nor understand the need for such a system. Congress, however, has proved itself a champion of scholarship and library preservation in the past, when it authorized and funded the brittle-books reformatting grants under the NEH. It acted then at the behest of scholars, who impressed upon it the irreplaceable value of the information printed on acidic paper. The task force recommends that a consortium of scholars and librarians, working with the AAS, the Library of Congress, and other appropriate bodies, develop a strategy for a series of repositories to assemble and preserve a full record of American publications. One such plan, for example, would have the AAS responsible for materials before 1876; another body or groups of institutions responsible for 1877 to the present; and the Library of Congress responsible for all prospective archiving. This archival repository system would be in addition to the local or regional repository system proposed above, which would serve the needs of scholars routinely requiring originals and would be designed to provide access to them. In contrast, an archival repository or series of repositories would be designed as a fail-safe mechanism that ensured the survival of an original, not ready access to one. The details of governance, deposit, access, and so forth would need to be carefully worked out by a number of interested parties, including librarians, preservation experts, and scholars.

3.3 Audiovisual

3.3.1 Sound and Light as Artifacts

Audiovisual materials present many of the same preservation challenges as do resources on paper. They are recorded on media that

decay and they are abundant, thereby forcing libraries and archives to make difficult choices in acquiring and preserving them. Researchers' demands for audiovisual products, like the demands for print resources, change over time. This reinforces libraries' sense that they need to collect masses of materials with no immediate demand in sight, in case they eventually prove to be of research value.

At the same time, in large part because of the extreme fragility of the recording media and the dependence on playback machinery that quickly becomes obsolete, audiovisual materials present new challenges to traditional notions of the intrinsic value of the artifact.[17] For printed materials, the artifact is a single physical object that has some measure of fixity. That fixity is largely dependent on the fact that the object has been recorded on a relatively stable medium. The stability of the paper depends on its physical composition, the effects of handling and storage on the original imprint, and other variables; however, there is little confusion about what a book or a journal is and what constitutes the original. Moreover, the role of the print artifact in research is fairly straightforward: it provides evidence, and does so through information that the physical object itself carries. This information offers a means of ensuring the authenticity of an object and of judging the veracity and accuracy of the information contained in it. It also brings researchers in some tangible way closer to the moment of creation of that object and, presumably, closer to the intentions of the creator.

A simple example can be used to illustrate the variety of ways in which audiovisual technologies assault the fundamental notion of artifact. Judged by any criteria, a film by Stanley Kubrick—*Spartacus*, for example—is important to keep and make accessible to future researchers, even if estimations of its merit fluctuate over time. What does it mean to preserve a film in its artifactual form? The Library of Congress, which acquired this film through copyright deposit shortly after its creation, defines the artifact as the original manifestation.[18]

The movie was a technological marvel during its day; however, we have now discovered that the 65-mm negative and 70-mm prints made from this negative both fade irreversibly, and sometimes dramatically, over time. What we have of "the original manifestation" of *Spartacus* is no longer a physical object that provides accurate or meaningful information about what viewers saw in theaters in 1960. The original object was fixed onto a carrier that precludes the very notion of stability. While there was a commercial effort recently to restore the film to a more or less faithful recreation of what it once

[17] "Intrinsic value is the archival term that is applied to permanently valuable records that have qualities and characteristics that make the records in their original physical form the only archivally acceptable form for preservation. Although all records in their original physical form have qualities and characteristics that would not be preserved in copies, records with intrinsic value have them to such a significant degree that the originals must be saved" (NARA 1982, 1). See also Menne-Haritz and Brübach 1999.

[18] The copy on deposit is likely to be in poor shape. Film companies routinely use exhibition copies, which are no longer fit for screening, to fulfill their legal deposit duty.

was, that restoration is not an original artifact from the 1960s, but one from the 1990s.[19] Looking ahead to the 2060s, there is every reason to anticipate that 35-mm and 70-mm films will be orphans of a technology that no longer exists, and that projectors to play them will be relics from the past. If no one reformats *Spartacus* onto new technology, then we will have lost direct access to a film that was crucial in the development of an important American artist who also created *Dr. Strangelove* and *2001: A Space Odyssey*.

This is a relatively uncomplicated example of what happens to the concepts of originality, fidelity, stability, and fixity when information moves onto one of the most significant recording media of the nineteenth and twentieth centuries. Similar examples abound for recorded musical performances and spoken-word documents.

While many humanists are familiar with print resources and the copyright regime that governs their distribution and permits timely preservation action, fewer scholars are familiar with the protocols that govern audiovisual resources. Their manufacture, preservation, and access protocols and, perhaps most significantly, the copyright laws that permit or restrict those protocols, are well-known only to specialists. Moreover, the history of collecting these materials differs dramatically from that of print resources. The following discussion of these matters, together with case studies in Section 4, provides background and a historical context with which to grasp the challenges that these media pose to scholars and librarians.

The curatorial and preservation communities involved in audiovisual materials are widely dispersed—in media, government, and public and research libraries, as well as in museums, historical societies, and regional collecting institutions. There are as many conflicts as there are common interests between the creative community (such as film directors, photographers, and musical artists) and the rights holders (such as studios, news services, and major media companies), or between researchers (folklorists, ethnographers, and linguists) and their subjects (the individuals, communities, and sovereign Native American nations who have been recorded). These conflicts sometimes have serious commercial consequences, and the communities can come to blows over preservation issues such as colorization and letter-boxing, cropping and retouching, or enhancing recorded performances for re-release on CD by cleaning up the noise in an original long-playing (LP) record. Because these resources often have high entertainment and commercial value, the role of the marketplace must be considered in any strategy to preserve and make them accessible for research. In other instances, for unpublished and noncommercial recordings, serious ethical issues have arisen between some researchers and the communities that have been documented. This has led to new access restrictions on old bodies of evi-

[19] Characteristically for the 1990s, the restorers tried to recreate what Kubrick wanted to make, not what the studio had released in 1960. For example, a scene with sexual innuendo between characters played by Laurence Olivier and Tony Curtis, which had not been included in the original cut but was restored in the 1990 version of the film, became a talking point in the publicity surrounding the restoration.

BEST COPY AVAILABLE

dence that beg the question of whether or not repositories should invest in preserving materials to which they may never be allowed to grant access. Some of the intricacies of rights management, commerce, and ethics are discussed in detail in two case studies about film preservation and folklore collections (see Sections 4.2 and 4.3). We focus here on describing the scope of visual and sound resources and highlighting some of the issues that affect artifactual evaluations and preservation options.

3.3.2 Still Images

The member libraries of ARL reported holding more than 64 million photographs, pictures, maps, prints, slides, charts, posters, cartoons, engravings, and other graphic arts in 1998–1999 (ARL 1999a). The Library of Congress has more than 13 million items in its photography and print collections. Many special-collections repositories also have significant still-image collections, often filed with printed, cartographic, or design materials in collections organized not by medium but by provenance or subject matter. No one knows exactly how much of this type of material exists in the United States or where it is stored. Nonetheless, it seems clear that the majority of historical documents that are image-based are not in academic research libraries but are scattered in historical societies, natural history collections, special-collections libraries, commercial archives, and local, municipal, state, and federal records offices. The charts that document land ownership and management; architectural drawings and engineering records of the built environment; photographs taken in the course of collecting news or creating journalistic essays; and archives of architectural firms, industrial design companies, and advertising enterprise—all these collections constitute invaluable historical records. While many companies take excellent care of their business archives and offer some level of preventive preservation for the still images scattered throughout their records, most enterprises do not have preservation strategies for their archives, lack resources to develop and implement such strategies, and may not even be aware that they hold materials that are of great potential value to historians and others. The task force recognizes that many collections that are part of the visual record of this nation lie well outside the purview of research institutions and that special efforts are required to identify and properly preserve these materials.

Within research institutions, still-image collections have been slowly coming into their own as primary source materials. In public libraries such as the New York Public and the Library of Congress, both of which have premier visual resource collections, most users are professional picture researchers who are looking for images to reproduce in publications. There is still a gap between those scholars who seek out visual sources as primary documentation and those who come to picture collections because they are looking for material to illustrate a monograph or an article. While professional picture researchers will always constitute a significant portion of users in public institutions, there is evidence that the current and future gen-

erations of humanities scholars will turn increasingly to visual documents for primary evidence.

There are several reasons for this trend in research methodology. The first is chronology, pure and simple. The primary recording media of the past 100 years have been visual, and anyone studying that period can—or in many cases must—rely on still or moving images for information. A second reason is that the proliferation of photography and easy reproduction of images in magazines and books have exposed the current generation to more imagery. Some believe that this has led to a rise in "visual literacy"; others argue that exposure to many images does not create the ability to "read," or understand, an image critically—the true sign of literacy. (Despite the proliferation of imagery in everyday life, few graduate programs outside the traditions of art history and archaeology teach students the same critical approach to visual sources that they do to textual sources. Reading documentary photography with an art historian's eye can be inappropriate for those pursuing other research agendas.) The third reason is that the topics and methodologies of literary, historical, and sociological research have broadened to include many phenomena that are not well documented in texts but are so in visual resources. Gender studies is a good example of a field that relies on a variety of sources from a variety of disciplines and makes good use of nontextual materials for such things as the history of domestic life (for example, illustrated magazines, advertising art, family photographs). There are also relatively new fields, such as environmental studies, that end up relying on the inadvertent documentation of built and unbuilt environments that would not have been remarked upon in texts. This priority of image over text is encouraged by the digitization of images. Visual resources perform superbly in an online environment.[20]

That said, what is the artifact in question? In the case of film-based images, is it the print, or is it the film from which the print was made? Sometimes the film has been developed and printed in a way that makes the printed image significantly different from the image captured on film. Most curators choose to collect both the medium of original capture and the print, when this is possible.

In pictorial collections, the items that receive priority for selection and preservation include any work that is of artistic value or of a certain age, or that has been created by a well-known artist photographer, graphic artist, or cartoonist; has a significant provenance; constitutes an institutional priority; and so forth. Original works—vintage photographs and their negatives or original drawings—receive the best care. Color film and photographs are generally more fragile than are black-and-white film and photographs; however, any photograph printed on resin-coated or acid paper is at risk.

[20] Libraries with important image collections have only recently begun to hire photo conservators. This is another indication of the increased importance of visual resources in research, although many conservators work not on materials that are in demand in the reading room but that are being prepared for digital conversion.

Film-based materials are often served in a high-quality surrogate, such as a digital image, slide, or copy print, unless the researcher is primarily interested in the object rather than the image content. A researcher using Toni Frissell photographs from the front of World War II could be satisfied with copy prints if she is looking for general-level information about Frissell's coverage of the war or various activities on the front. Another researcher might be interested in seeing only vintage prints, negatives, and contact sheets that carry information about the entire project or shooting assignment, the original sequencing of shots, how the negative was cropped and printed, and so forth.

One of the dilemmas of preservation is that the best conditions for storing an object often compromise the object's effectiveness for research. Still images, whether film or paper based, are often an integral part of a mixed-media collection. Such images might include photos in a personal archive, illustrations in a book, or architectural records in a business archive. Preservation would demand that they be physically separated from the original collection and stored elsewhere in cold vaults. Each medium has its own storage requirements; for example, requirements for black-and-white film differ from those for color.

There are large collections of documentary photography that gain research value as they are supplemented by a range of secondary materials—notebooks, work files, and contact sheets—that enlarge the evidential base of the documents under review. In many cases, it is sufficient to create surrogates that can be kept with the original collection. Copy photos or digital reproductions usually provide all the information needed, but creating high-quality surrogates can be costly.

3.3.3 Moving Images

The catastrophic loss of silent film—the incunabula of moving images—that occurred before movies came to be seen as respectable sources for research should warn us of the perils of undervaluing new media. What we have faced in film is not only a failure to keep films that have been collected from undue deterioration but also a failure to collect films systematically at all. Fortunately, certain events conspired in the last 15 years to draw attention to the lamentable state of film preservation and galvanize communities into action. This should serve as an example of what can and should be done for other audiovisual sources.

While both public and private nonprofit archives have long been working to document and preserve film, it was the emergence of ancillary markets for resale of film inventories that moved the industry itself to invest resources in preservation. As videotape playback equipment became cheap enough for the consumer market, studios predicted that recycling old films for reissue would be worth the investment. That led to the introduction of colorization, theoretically to enhance the appeal of old movies to consumers. Colorization and cropping to fit a television screen, in turn, galvanized the creative

community around issues of original intent and moral rights, which sparked a confrontation with commercial forces and a widening circle of discourse about the state of the global film heritage.

In 1988, Congress passed the National Film Preservation Act, creating the National Film Preservation Board and asking the Copyright Office to investigate the colorization and material alteration to film. By the time the film board was reauthorized in 1992, its focus had moved from alteration to preservation in the broadest sense. A consequence of its work was that all the players—archives and libraries, studios, artists, and distributors, each with competing interests—were forced to identify their common interest in the integrity of the film heritage and to collaborate for the first time toward a common preservation goal. This concern for the integrity of and continued access to film was one factor in the development of a national strategy for preservation. It coalesced easily with other significant initiatives of the time, notably the formal establishment of the Association of Moving Image Archivists (AMIA), which brought together the nonprofit and for-profit professionals in film—archivists, technicians, filmmakers, academics, and laboratory managers. The high-profile efforts of filmmaker Martin Scorsese and the Film Foundation also raised awareness of these issues among the public and creative communities. In 1993, the film board undertook a national study of preservation needs and the following year put forth a national plan (LC 1994). In 1997, with congressional authorization to fund film preservation, the National Film Preservation Foundation opened its doors with the financial support of the Film Foundation and the Academy of Motion Picture Arts and Sciences.

That landmark film-preservation study addressed a broad range of issues, including technical, legal, economic, and financial issues around preservation. It helped document what had been created, what had been lost, and what was at highest risk for loss. The first order of business was to document what has been lost, and that was tricky. Written records, old journals and newspapers, and collections of film memorabilia provided clues. A mental reconstruction of the film industry at certain periods of time led to meaningful deductions about what must have been and to an understanding of the distribution chain that could suggest where copies of films had been shown. This, in turn, led to happily correct suppositions about where missing films could be located. The American film industry often shipped abroad films that, for various reasons, were never returned at the end of a run. Australia and the Czech Republic proved to hold significant American films thought to be long lost. An international collaboration is crucial to the ongoing reconstruction of the film record of any nation.

Documentation is a serious challenge for the film community, especially for noncommercial films. Because so much of what has been created is either lost or in imminent peril of disintegration, information about what once was made and how it was distributed and received can serve very important needs, even without the original itself. Film, after all, is a public medium with wide influence on

those who see it. One example of a documentation project on which scholars and archivists have collaborated is the British Universities Film & Video Council (BUFVC). It has created a database to document what newsreels were shot and shown in theaters, even those newsreels that no longer exist.

From these state-of-film data, the National Film Preservation Board and its collaborators developed a strategy for preserving films that was based on a clear division of labor. Studio-produced and -distributed films are preserved by a series of collaborative and unilateral arrangements. Because studios now see their old films as assets, they are preserving the artifact and its economic value. Studios are building new storage facilities. But while many studios preserve a significant portion of their inventory, a very large portion of films are actually preserved and stored at the four major film archives (the Library of Congress, UCLA, the George Eastman House, and the Museum of Modern Art), often with financial assistance from the studios.

The National Film Preservation Board's study also highlighted an array of noncommercial films, from independent art films to documentaries and home movies, that are vital parts of the historical record. Some "home movies," for example, were shot at an internment camp for Japanese Americans. Because these films and their importance have now been identified, public and private support for funding their preservation has emerged. The National Film Preservation Foundation was established to coordinate these preservation efforts. Its approach is to find funding from a mix of private and public sources to regrant to those repositories—historical societies, regional film archives, local museums—for the purpose of undertaking preservation work.

The preservation challenges are complex, and the solutions are expensive. Film-based materials need to be protected not only from their inherent chemical instability but also from the mechanical damage incurred by running a reel through a projector. A master negative needs to be properly stored and used only to produce intermediaries that can be used for creating access copies (either film or tape). Original materials for old films, no matter how badly deteriorated, need to be saved in the event that a new technology comes along that can extract more information from them. The four major film archives in the United States all keep original materials.

There is much dissension among preservation experts about whether analog materials should or should not be preserved in digital form. With respect to film, the important thing to note is that the physical artifact itself is an endangered species that warrants special measures to ensure its survival. Few institutions have taken on the burden of preserving film. The best practices that have emerged include retaining as much of the original material as possible, including any supporting materials that provide further evidence of the film's original state or that can be used to restore lost portions of films. For the sake of preserving the original materials, researchers are given access to reformatted versions. Technology is increasing the

fidelity of reformatting so successfully that most researchers do not need access to the original.

3.3.4 Recorded Sound

The heritage of recorded sound is imperiled because the industry that manufactures currently acceptable preservation media—primarily analog reel-to-reel tape—is phasing them out, along with the playback equipment needed to render the encoded information intelligible. Sound recordings seem to be especially vulnerable to technological obsolescence. Anyone who grew up with 78s, 45s, and LPs, then cassettes and CDs, and now reads about (or downloads files through) Napster and Gnutella, is aware of this shifting landscape. Nonetheless, these consumer formats have remained remarkably stable relative to the industry standards with which the preservation community has to keep pace. One expert reported the following to the Library of Congress:

> Most industry representatives report that their audio and video products—media and systems—will all be digitally based within the next five years or less. This abandonment of traditional analog technology has come alarmingly fast and in dramatic fashion. Access, even on a limited basis, will need to become digital. Master preservation copies will need to become digital (Storm 1998, vi).

Many people disagree about whether digital preservation of film is necessary for analog films, because commercial entities still shoot with old-fashioned film stock. Editing is now done primarily on digital equipment, and some major filmmakers—George Lucas most famously—are shooting with digital cameras. As long as film stock continues to be manufactured, the debate will go on.

Recorded sound is a different matter. In this case, members of the preservation community will soon find themselves without the tape and equipment they need for analog reformatting. Acquiring the equipment needed to transfer analog to digital is a huge capital expenditure—one that many institutions that hold valuable recordings are not prepared to make. Consequently, many collecting institutions will have to outsource their preservation reformatting. Such a process is fraught with risk because few laboratories specialize in the art of old media.

Among the most significant repositories of sound recordings are the Library of Congress, the National Archives, the New York Public Library, and a number of academic institutions, including Columbia, Yale, Syracuse, and Stanford Universities. There are also important, but smaller, collections such as those of whale sounds at Scripps Oceanic Institute and Cornell University's ornithology collections. Folklore and linguistic materials abound in the personal collections of many scholars. Some type of preservation needs to be done for these items as well, because they constitute primary data for these scholars' research. If one of these personal, unpublished, and often uncataloged, collections goes, the evidence base for that scholar's work

also disappears. It is fair to say that for field recordings that exist on various tapes and cassettes, the artifact itself has little to no value; the chief mission of preservation is to rescue the information on those media before it disappears. It is crucial that scholars attend to their own collections, and they that they do so now.

Perhaps the most important thing that scholars can do to prevent the loss of these documents is to recognize their value and to lobby for support for their preservation. The U.S. Congress, largely in response to pressure from constituents and from the industry, recently passed the National Recording Preservation Act of 2000. Modeled on a similar act for film preservation, it calls for public recognition of the value of the nation's sound heritage and for a survey to document the state of recorded sound. (See Appendix VII for the full law.) Until that survey is completed, it will be hard to assess accurately what is imperiled.

A significant part of the identification problem is that, outside the field of classical music, sound recordings have not been incorporated into the canon of bona fide research resources until recently. The copyright regime has been slow to grapple with the fact of recording. Recorded sound has been covered by the copyright code only since 1972. Even when one admits popular and ethnographic music into the fold, libraries have collected chiefly published recordings. Published materials have not been readily incorporated into the bibliographical systems of libraries, and unpublished materials have almost never been included in these systems. Researchers have found that commercial catalogs are the best sources for the initial phases of a search of published recordings. For unpublished materials, which constitute the core collections of most anthropologists, ethnomusicologists, folklorists, and linguists, few if any points of access are available. Scholars who are themselves creators of documentation should start thinking about access and preservation issues at the moment they create the documentation, from getting informed consent from subjects so that rights to access for various purposes will not later become a problem, to using best practices for recording, describing, and storing materials.[21]

At present, selection of audio for preservation, like that of moving image material, is based on an assessment of the following factors:

- cultural value of the item
- historical uniqueness of the item
- estimated longevity of the medium
- current condition of the item
- state of playback equipment
- access restrictions
- frequency of use

[21] For a comprehensive view of audio preservation and best practices for recording, see the Web site of the Photographic and Recording Media Committee of the Preservation and Reformatting Section of the American Library Association. Available at http://palimpsest.stanford.edu/bytopic/audio/.

When the access copy or preservation copy does not adequately capture the information, it must be retained for future use because of the probability of advances in technology.

3.3.5 Broadcast Media

Broadcast media—radio and television—are the stepchildren of the audiovisual media. While everyone acknowledges their reach across the country and their power to shape public perceptions, interests, and thinking, few institutions take them seriously as sources of historical information, at least as evidenced by how few collect them. Vanderbilt University collects television and specializes in network news. The Library of Congress has an unparalleled collection of black-and-white-era television holdings and some significant radio holdings, such as the Armed Forces Radio and Television Service Collection. But in general, at both the national and the local levels, what gets broadcast is woefully under-documented.

There is little local collecting, yet because broadcast media are the equivalent of newspapers in another era, the most significant materials are generated locally and should be collected and preserved at the local, not national, level. Those who decry the loss of newspapers from the past are well positioned to raise awareness about the present need to save broadcasting, because what they most value about old newspapers is present in the newer forms of news and entertainment—exquisite sources for popular opinion, advertising, and cultural phenomena of all types.

The National Historical Records and Publications Commission (NHPRC), a body of the National Archives, has recognized the problem and become a leading funder of local television news collection. The Library of Congress has been authorized to create an archival record of national radio and television through the American Television and Radio Archives program. However, the state of local television news collections across the country is, in the words of an expert report, "extremely desperate" (LC 1997, 91). The state of collections of non-news television and of radio of all sorts is equally deplorable, if not more so.

The Library of Congress report on the state of television and video preservation (1997) lays out the scope of the problem and proposes a national plan of shared responsibilities to ensure that at least some of the television and videotape in the country will survive. Should these broadcast documents not survive, it would be an irreparable loss for present and future generations. Scholars can play a crucial role in making the academic communities of which they are members, as well as the public, aware of the value of this documentation, the extent of the challenges to future access, and the efforts under way to preserve these materials. It would be ironic and sad indeed if this generation of scholars, using ephemera and other artifacts of the nineteenth century to reconstruct the history and consciousness of that time, were to do nothing to articulate the need to safeguard the comparable artifacts of this era.

To the extent that preservation funding follows closely what the scholarly community has declared to be of research value, all the audiovisual formats discussed in this report suffer at present from lack of advocacy by scholars. But the task force recommendations cannot assume that significant additional funds will be forthcoming simply because the need for them has been identified. University librarians and administrators alike have testified to the task force that preservation budgets today are remaining flat or, in real dollars, shrinking. There is no reason to believe that this trend will not continue. If some base funding were to become available for reallocation, it would surely go to meet the greatest faculty need—the purchase and licensing of electronic resources. However, in the area of film, television, recorded sound, videotape, and other audiovisual media, partnerships beyond the academy are critically important, and the national preservation plans that are in place, or will be so shortly, must claim the time and attention of the scholarly community.

3.4 Digital

At first, digital information objects appear to be outside the scope of work of the task force, charged as it was to investigate the role of the artifact in library collections. After all, what is an artifact if not a physical object? And what is digital information if not intangible, available to be "output" to any number of media for access but having no intrinsic physical form and not bound by the temporal and spatial constraints of print, visual, and sound resources? Is it not digital information itself that is threatening to replace, or at least displace, the physical collections that libraries hold?

It is in part the phenomenal growth of digital information in libraries that prompts this examination of the nature and importance of physical artifacts for research and teaching. Digital technology is changing the ways in which researchers and teachers are getting access to information and in some cases obviating the need to consult an original. The demand for electronic resources and the infrastructure needed to support their access and maintenance is affecting the budgetary climate in which artifactual preservation must fight for its share. Finally, the flooding of the information landscape by "nonartifactual" intellectual property influences our understanding of the notion of the artifact in ways that beg scrutiny. It is for reasons such as these that digital formats warrant special attention.

Thus far, the task force has focused primarily on the value of digital conversion for the purpose of access to originals—on the creation of surrogates of nondigital works to overcome obstacles created by scarcity or physical fragility. With respect to original artifacts, the salient questions to pose about digitized representations are as follows:

- When can a digital surrogate stand in for its source?
- When can a digital surrogate replace its source?
- When might a digital surrogate be superior to its source?
- What is the cost of producing and maintaining digital surrogates?
- What risks do digital surrogates pose?

BEST COPY AVAILABLE

In print, there is a great deal of secondary literature that is amenable to digitization. There is also a great deal of primary source material—dime novels, vintage photographs, and various ephemera—that, because of its scarcity, could be more widely accessible were it in digital form. Other items—for example, oversize materials such as posters, or books in fragile states—are easier and safer to use in the form of digital surrogates. The research tools available for digital materials, such as full-text searching, may make the surrogates more accessible to specific types of research questions than the originals are. Postscanning processes may make the use of damaged items easier because one can lighten dark patches or sharpen the resolution of faded inks.

For visual resources and recorded sound, reformatting for access purposes is nearly always recommended or required for the sake of preserving the integrity of the source material. There is a growing consensus that digital, rather than analog, reformatting will best meet the demand for accessibility, fidelity, ease of reproduction, and cost-effectiveness, although significant issues still must be addressed in standards for access, preservation, and rights management before this technology can fulfill its promise. For audiovisual resources, the funding of the preservation work that is required remains a hurdle.

In sum, there are many ways in which this new technology can create adequate and at times superior access to information in physical artifacts. There are also instances in which no surrogate, no matter how splendid, will serve the scholar's needs. Finally, the infrastructure needed to support and sustain such reformatting programs is still in its infancy.

Before one can recommend digital reformatting for the preservation of or access to artifacts with intrinsic value as physical objects, it will be necessary to identify the hazards of digital representation of artifacts and to determine the true nature of an artifact in a digital environment. Physical artifacts bear all sorts of evidence about how they were created or manufactured, who had possession of them at what time, how they were used, how they have changed, whether the information they contain has been altered, and whether the alteration was intentional or inadvertent. What happens to this evidence when the object is represented in digital form?

3.4.1 Artifacts and Artifactual Value in the Digital Realm

Libraries and archives have two distinct categories of digital objects—materials that exist only in digital form ("born-digital" information) and materials that are digitized versions of analog source materials ("reborn-digital" information). The two categories can exist side by side on a server, and the information technologists responsible for the maintenance of those files do not draw distinctions between them in terms of treatments. From the researcher's point of view, however, the distinctions are great, and they will be maintained here.

Digital technology can represent all genres and types of library and archival materials—textual, numeric, visual, and sound informa-

tion. Because the technology of creating and disseminating information through digital means is relatively new, society has appropriated many terms specific to analog information and used them, almost by analogy, to describe digital objects. This goes beyond such examples as referring to individual instances of files uploaded to computer screens as "pages." It has also resulted in such awkward back-formations as "nondigital" information to mean everything that already exists in analog form.[22] These neologisms and appropriations have considerably complicated the work of the task force, and none has proved more vexing than the use of the word "artifact" in the digital realm.

The simplest, and one of the oldest, denotations of the word "artifact," as discussed previously, is a physical object on which information is recorded. The very value of the physical manifestation, together with the dependence on the physical medium, has created problems for preserving information printed on paper. That problem deepens with audiovisual information objects, because the mechanical processes that produce and record audiovisual information produce physical artifacts far more fragile than paper is. The concept of artifactual value is contingent not only on changes in cultural valuations but also on information technologies and their businesses—be they printing presses and paper mills or film stock manufacturers and playback equipment companies. All these complexities abound, and are perhaps even more problematic, in digital formats as well.

Some task force members argued correctly that an artifact is something—anything—made by art and that it does not need to be physical. This is one of several definitions the *Oxford English Dictionary* attests to the word. Nonetheless, the notion of physicality was intrinsic to the working definition used by the task force, because its central charge was to distinguish between those times when the researcher needs the original physical manifestation and those when a secondary or reformatted manifestation is sufficient. This is yet another example of how slippery a concept "artifact" is at heart. To speak of a "digital artifact" may even appear paradoxical, to the extent that "artifact" implies uniqueness or scarcity, whereas electronic information is replicable to the point where "original" and "copy" may lose their frame of reference. Consequently, when considering artifacts that are originally digital, the first and possibly the most difficult question is "What is the artifact?" This question is discussed in detail in Section 2. We will return to it during the discussion of born-digital items.

3.4.2 Digital Surrogates

What is the utility of a digital surrogate? The answer to this question depends, to a large extent, on the nature of the original artifact and

[22] Though inelegant, the word "analog" is preferred in this report to "nondigital," because in the present context, the identity of these (analog) collections is not derivative of their distinction from digital formatting. "Analog collections" is also preferable to "legacy collections," which has the misleading connotation that all future collections will come only in digital format.

the conditions of its use. Therefore, as a means of determining the value and appropriate use of digital surrogates for library holdings, it may be useful to divide original materials into those that are rare and those that are not, and to divide them further into those that are frequently used and those that are infrequently used. There would be, then, four possible cases:

1. Materials that are *not rare* and that are *frequently* used. In this case, we can assume that preservation of the original is not a particularly high priority (since the original is not rare); nevertheless, digital surrogates for such an object might be worth producing and providing, for several reasons:

- to reduce the cost associated with reshelving the object
- to make the object simultaneously available to multiple users (for example, through an electronic reserve desk)
- to replace the object, thereby doing away with the cost of housing it

The first two are obvious and uncontroversial benefits. The third is potentially problematic, even if the object in question is not rare, because it is not obvious that digital surrogates provide all the functionality, all the information, or all the aesthetic value of originals. Therefore, while it may be sensible to recommend that digital surrogates be used to reduce the cost and increase the availability of library holdings that circulate frequently, the decision to deaccession a physical object in library collections and replace it with a digital surrogate should be based on a careful assessment of the way in which library patrons use the original object or objects of its kind. It is not necessary that the digital surrogate possess all the qualities and perform all the functions of the original, but it is necessary that the digital surrogate answer to the identifiable needs and expectations of those who frequently used the original.

2. Materials that are *not rare* and that are *infrequently* used. Many libraries now store infrequently used books and other materials in long-term storage facilities. Those materials are retrievable and available to library patrons, but only after a wait of two or three days. With such materials, digital surrogates might

- help users to determine whether recalling an object from long-term storage was worth the wait—and worth the library's effort
- increase frequency of use (by providing searchable metadata, for example)
- reduce costs by replacing the object with a digital surrogate

The first two are obvious and uncontroversial benefits, and the third comes with the caveat that the digital surrogate should answer to the identifiable needs and expectations of those who (in)frequently used the original. At some point, especially with infrequently used materials that are not rare, libraries might be expected to evolve a calculus that balances functionality with actual use in order to help decide when digital surrogates that provide *most* of the functionality of originals are acceptable.

One other point needs to be raised, especially here, where we are discussing the component of library collections that has the least

BEST COPY AVAILABLE

"market value." Libraries, as an institutional and cultural community, need to consider whether these infrequently used and commonly held materials are, in fact, being preserved in a concerted and deliberate way in their original form by any one (or more than one) library. If they are not, the sources for digital surrogates that are common today could easily become rare, or nonexistent, tomorrow. This is the substance of Nicholson Baker's objection to libraries' practice of discarding their newspaper holdings. If 50 libraries are holding the same issues of the same newspapers in original form, at great expense and with limited use, then it is difficult to make the case that all of them should pay to house, shelve, reshelve, and preserve the originals. However, if 49 of those libraries, over time, have replaced their physical holdings with digital surrogates, one certainly hopes that the fiftieth library would be aware that its physical holdings were now rare, and therefore subject to considerations outlined directly below under cases 3 and 4.

3. Materials that are *rare* and are *frequently* used. In this case, the principal (and very obvious) benefits of digital surrogates are

- Preservation: By standing in for some uses, the digital surrogate reduces wear and tear on the original object; and
- Access: By providing access that does not impose wear and tear on the original, the digital surrogate makes rare objects more accessible.

Few would argue that digital surrogates should replace truly rare materials. Digital technology and techniques of digitization are so new, and are developing so rapidly, that we cannot be confident that we have devised the best method for extracting and digitally representing information from any analog source—whether it is a printed page, an audiotape, or a filmstrip. Nonetheless, digital surrogates could, in many cases, stand in for rare and frequently used materials, and could thereby aid in the preservation of originals.

4. Materials that are *rare* and are *infrequently* used. On the face of it, these materials seem the least likely to be represented with digital surrogates, if only because digitizing is expensive. On the other hand, if the cost of housing a rare but infrequently used object rises high enough, then digitizing and deaccessioning that object may become an attractive possibility. Here again, libraries need to be aware of the actual or potential rarity of even those materials used infrequently today. Tomorrow, those may very well be the most valuable of artifacts, perhaps for users, or uses, that one could not predict today.

The basic questions, and their answers, are therefore as follows:

- When can a digital surrogate stand in for its source? When it answers to the needs of users.
- When can a digital surrogate replace its source? When the source is not rare.
- When might a digital surrogate be superior to its source? In cases where remote or simultaneous access to the object is required, when software provides tools that allow something more than or different from physical examination, and when the record of the

digital surrogate finds its way into indexes and search engines that would never find the physical original.

- What is the cost of producing and maintaining digital surrogates? The cost of producing digital surrogates depends on the uniformity, disposability, and legibility of the original. The cost of maintenance depends on frequency of use and the idiosyncrasy of format. Beyond that, cost depends on technological, social, and institutional factors that are difficult or impossible to predict. This is an important reason for being cautious when one chooses to *replace* a physical object (the maintenance costs for which are known) with a digital surrogate (the maintenance costs for which are, to some extent, unknown).
- What risks do digital surrogates pose? The principal risk is the possibility of disposing of an imperfectly represented original because one believes the digital surrogate to be a perfect substitute for it. Digital surrogates also pose the risk of providing a partial view (of an object) that seems to be complete, and the risk of decontextualization, i.e., the possibility that the digital surrogate will become detached from some context that is important to understanding what it is and that it will be received and understood in the absence of that context.

3.4.3 Access

While digital collections such as the William Blake Archive or the Women Writers Online databases have been created (in part) to make disparate items more accessible, scanning is but the first step. Ease of access depends on how these items are described in cataloging and index schemes and how easy it is to find and retrieve records of these collections in a catalog or database. The cost of cataloging, mark-up, and other things that make up metadata (i.e., data about data) is often as high as, or even higher than, the cost of image capture. This cost is also one of the few things about digitization that is not likely to go down as a result of technological improvements, because it requires significant human input. It is important that descriptions of digitized materials be done in the most accessible forms possible, not in hand-created systems or forms that are commercially restricted. Creating records for monographs and serial publications can be fairly straightforward and can allow direct linking to an institution's online catalog. Visual resources can be marked up according to the controlled vocabularies found in *Art & Architecture Thesaurus* and *Thesaurus for Graphic Materials*.

Fields that do not have a controlled and shared vocabulary—folklore, for example—are at a great disadvantage in a networked environment. Many ethnographic and field recordings exist in analog formats that are rapidly deteriorating or becoming obsolete or that have a limited life span. They are prime candidates for digital reformatting. Saving the information by transferring it from audiocassette to digital files is only the first step. The files need to be described in conventional language to allow access. It is no accident that textual corpora have been among the first to find new lives in

digital form. These are materials that have standardized forms and bibliographical standards. The machine-readable catalog (MARC) record is a standard that allows libraries to ensure uniformity of description and to share information in a networked environment. Visual and sound resources, as well as manuscript collections, are less standardized in description and so are rendered less readily accessible online. Librarians and archivists are making efforts to devise descriptive systems that allow for uniform descriptions of archival collections; Encoded Archival Description (EAD) is among the most widely adopted system so far. Moreover, key professional associations in sound and moving image archives are developing best practices for digital capture and mark-up. Academic fields such as folklore, dance history, and ecology, which lack best practices in documentation and research, would be well advised to agree on ways to make their resources accessible through standardized descriptive practices. Access tools will be user-friendly only to the extent that users are involved in their creation. It is time to engage systematically in studies of research needs and researcher behaviors online, so that the tool kit developed for their use will meet the users' needs.

Another missing feature of the current digital library infrastructure is a central location from which to discover which analog collections have been digitized and how one can get access to them. There is much talk in the digital library community about building a registry or series of registries for digitized items, but so far, none exists (*RLG DigiNews* 2000, DLF 2001b, Greenstein 2001). One reason is that organizations that digitize use a great variety of standards, not only for capture but also for description. Some items are described at the item level, others at the collection level. Items that have been digitized more than once would not all appear in a search of the records as they currently exist. Libraries are often short-staffed, and they cannot report regularly on what they are doing or planning to do. The absence of such reports would jeopardize the timeliness and reliability of such a registry. No single body has agreed to act as an organizing agency to create and maintain the information, and it is not clear from where the funding for such a body would come. The work of such a body would include securing the information and ensuring that changes in URLs were kept up-to-date. In the meantime, a number of associations and groups at the local, state, and national levels have created databases about what they are doing. These, taken together, do not meet the researcher's need to locate without great difficulty the digital surrogates of items he or she seeks.

3.4.4 Born-Digital Materials

Whereas digital surrogates always originate from physical source materials, a born-digital item has no previous manifestation in physical form. It is entirely dependent on hardware and software for accessibility, storage, and long-term access. Anyone who has tried to trace a citation to a digital source, only to find that the site no longer exists, understands that dedicated maintenance and resources are

required to keep digital sources alive, let alone up-to-date. Beyond the problem of stability, it can be difficult to judge the authority or authenticity of Web-based information.

Digital information is by its nature perfectly replicable. To distinguish between the first and the forty-first instantiations of a digital file is a fool's errand. But born-digital information is of very great import for scholars interested in the artifact, for it challenges notions of originality and uniqueness, and even of authenticity, fixity, and stability. That it does so matters greatly, because of the sheer quantity of information that is being created in digital-only form. In 1998–1999, ARL libraries spent an average of $742,598, or a median of 10.18 percent of their materials budgets, on electronic resources (Jewell 2001). While still a small portion of what is spent on total acquisitions, this percentage is increasing each year. This category of materials would include databases such as the Inter-University Consortium for Political and Social Research data sets (the world's largest archives of computerized social science data) or the Environmental Systems Research Institute data and maps (a detailed set of data for the United States and the world that includes census boundaries and major transportation features), and—increasingly—thematic research collections. Such examples as Ed Ayers's "Valley of the Shadow" Civil War site (Ayers et al. 2001) and Columbia International Affairs Online, while largely based on digital surrogates, also include layer upon layer of secondary or editorial work that is uniquely digital, and the sum of the parts is wholly digital. Many of these new "digital objects," as they are often called, have been created by scholars or publishers or both. They have the same claim to intrinsic value that any intellectual property created with analog technologies would have. By far the largest portion of these electronic sources, however, is the science, technology, and medicine periodical literature, an increasing percentage of which is available exclusively in digital form.

Of the things we value in these materials, what is fungible and what cannot be separated from its carrier medium? How does one begin to make that distinction? In many digital objects, the chief value is in the informational content of the file, and that content is as good in one copy as in another. It is similar to the case of feature films, where the researcher will normally prize ease of access over fidelity to the original manifestation, within limits. Many researchers are happy enough with laser-disc or VHS-tape versions of films they are researching. So it is with digital information: ease of access is usually its greatest value. A PC Word document viewed on a Mac will still serve the access needs of the researcher. It is seldom important to seek out a PC to view that same document with full fidelity to the look and feel of the environment in which it was created.

But are there equivalents in the digital realm to the intrinsic values we have defined for the physical artifact, namely, originality, faithfulness, fixity, and stability? This is a subject of intense discussion among librarians, archivists, computer scientists, and security experts. A detailed examination of these complex issues is beyond the scope of this report. A summary of the major issues surrounding

the problem of authenticity is, however, necessary for understanding the possible hazards of using digital information.

Originality in its simplest sense—that is, of "oneness"—is not a meaningful notion in the digital realm. Archivists have many ways of defining "digital original," all of which may suffice to meet core requirements for evidence. In a library context, however, the notion is more or less bankrupt (Bearman and Trant 1998, CLIR 2000, Lynch 1999). Faithfulness, by contrast, is not necessarily an empty notion. It is true that the bit streams that make up a file can be changed without evidence of the change, regardless of whether it was intentional or accidental. For example, someone who receives a forwarded plain-text file that has had three paragraphs deleted from it would not be able to see evidence of the deletion on a computer monitor. Internal evidence, however, might lead an attentive reader to believe that "something is missing," and if the file comes with a digital signature—a public key or other mechanism for objectively verifying that the file has not been changed since the key was generated—then one does have some hope of determining faithfulness.[23] Of course, few researchers perform forensic examinations on the physical artifacts they view. Users trust that a source found in a library, for example, is more reliable than one found posted to a billboard. So, too, many researchers place more trust in resources found on a library Web site than in information anonymously posted on the Web.

The issues of stability and fixity, however, are quite troublesome for digital texts. For analog sources, stability and concrete form are absolute requirements for proving the authenticity and provenance of unique documents. The malleability of digital information, the ease of creating several different and subtly tailored versions of digital documents for different audiences, and the difficulties of maintaining digital sources intact and accessible over time are similar to the problems associated with archiving the international television broadcasts cited in the previous section. Which version of a digital document designed to evolve and grow over time is the one that should be archived? For interactive documents, what importance does the interaction have in defining the authentic document? It is important that scholarly societies work with publishers and librarians to specify what should constitute the archived version of an e-journal article, for example, and who should be responsible for creating that final version. Efforts are under way to bring together libraries and publishing houses to address the many questions that surround the archiving of scholarly journals (DLF 2001c). It is in the interest of members of the academic community to join this effort.

As to the intrinsic value of the digital artifact, as opposed to the value of the content that might easily be reformatted onto another medium, what are those features that drop out when content is reformatted? It could be the text formatting, as anyone who has imported

[23] See the application of digital authentication technology to electronic scholarly editions, carried out by the Australian Scholarly Editions Centre. Available at: http://idun.itsc.adfa.edu.au/ASEC/PWB_REPORT/Review.html.

BEST COPY AVAILABLE

a file created in one word-processing program into another program well knows. But in most cases, it is some sort of functionality that is lost. Relational databases can export their data in a form that permits another database engine to import that content, but the relationships laid down among items in the original database will be lost, because there is no standard way of encoding those relationships. Without those relationships, much of the functionality of the database will be lost as well. So, just as with analog sources, preservation of the digital artifact is successful to the degree that it maintains over time the chief distinctions of that object as digital—its functionality, its formatting, or whatever is important about the digital object for a particular use and a particular user.

3.4.5 Preservation of Digital Information

Computer scientists, librarians, and archivists have explored the subject of preserving digital information in depth for nearly a decade, and it continues to be a subject of active research. This section summarizes a topic that has been well covered by others.[24]

Like analog moving-image and recorded sound information, digital information is both technology- and medium-dependent. Digital technology appears to be evolving even more rapidly than analog recording media have evolved. Hardware and software have changed so quickly that information that has been recorded on one software program often is at risk of obsolescence in only three or four years. The media on which digital information is recorded, whether it is magnetic tape or a hard disk, are vulnerable to a startling array of new risk factors, including magnetic fluctuations, separation of information from the substrate caused by mild environmental disturbances, accidental overwrites, and "bit rot."

Precisely because the technologies used to encode, display, and enact digital information are changing so rapidly, the digital artifact that goes untouched for 10 or 20 years may well be unrecoverable. Its storage medium might require hardware that no longer exists, the software used to create it might no longer be available, or the operating system under which that software ran might be obsolete. Digital files must be "refreshed" and magnetic tape "exercised" to ensure that the bits keep their integrity and the recording medium does not suffer undue degradation. This constitutes minimal-level preservation for digital files. There is no such thing as putting the text on eight-inch floppies, putting the floppies in storage, and being able to retrieve the text 50 years later. Digital preservation programs will have to be much more active, and even aggressive, than book preservation programs are.

In the past, with information stored in more or less durable physical objects, the task of preservation has been to stabilize and, if possible, enhance the integrity of items (books, films, sound recordings) that had been around for generations because they had been

[24] See, for example, Task Force on Archiving of Digital Information 1996, Rothenberg 1995, Rothenberg 1999, Bearman 1999, and Granger 2000.

acquired by a library. Some judgment has been made about the intellectual or cultural value of the item, and the question for preservation is not whether, but how, to preserve the item. With digital information, whose life span can be as short as one software upgrade, the decision to preserve must be made almost simultaneously with its creation. This turns the traditional preservation paradigm on its head.

The two most commonly advocated methods of preservation, both of which involve concerted efforts by custodians, are *migration* and *emulation*. Migration is essentially a moving of digital files from old hardware and software platforms onto new ones. Migration is a translation, and some measure of loss is inherent in the process. What happens when one reformats word-processing files from an old version of a program to a new one, or from a Mac platform to a PC? The content of the documents is usually preserved, but the formatting is corrupted to one degree or another. Migration has occurred, and loss has been incurred. Yet most people who look back at the costs and benefits of the process would probably decide that the benefits of not having to rekey whole documents outweigh the tinkering necessary to restore the original formatting.

At the same time, there are categories of digital documents, most of them nontextual, that might suffer unacceptable losses during migration. These include executable files, such as time-lapse simulations and interactive programs. For these, the goal is to preserve not only the content but all aspects of the behavior of the original content and of the software used to present it. This is what emulation is designed to do—create software that can simulate the hardware and software environment in which a document was created. The technology of emulation is still in its early stages.

When considering artifacts that are born digital, the first and possibly the most difficult question is "What is the artifact?" What information or value inheres in the carrier medium? Is the equipment originally used for display part of the digital artifact? Does the software that presents and actualizes the data qualify as a constituent element of the artifact? Thinking again of the criteria for determining intrinsic value in an artifact—as a way of understanding what the features of an artifact might be—it is evident that there are a number of practical ways in which these questions might surface.

The physical form of a document or program might be the subject for study if the records provide meaningful documentation or significant examples of the form. For example, the layout of a form used for collecting data on the Web might reveal a good deal about otherwise-inexplicable aberrations, omissions, or misconstructions in the data collected.

There are also aesthetic or artistic qualities that may require preserving, most notably digital art and literature (Lyman and Kahle 1998). Achieving that level of fidelity to the original, or original instantiation, would require running a program on originally specified system hardware and software. Early computer games are a popular species of programming wizardry that inspire heroic efforts at emu-

lation, or recreation of original hardware and software platforms, for replay. One such emulation program can be found in Java, which runs Kaypro (Z-80/CPM-based) software on contemporary Intel/Windows-based systems. The challenges posed by preserving digital art are manifold, because the art is often designed, like a game, to provide real-time experience for the viewer—an experience that may or may not be intended to be replicable. The Guggenheim Museum of Art has developed a program, Variable Media Initiative, that "encourages digital artists to help establish preservation guidelines long before their equipment, code, and lives become history," as *Wired* recently reported (Jana 2001). Through this program, artists can record not only the technical specifications of a given work of art but also the artist's intentions or considerations for longevity.

A third form of digital preservation is the preservation of the original hardware and software for use at some time in the indefinite future. Similar to the saving of wax cylinders and the original play back equipment for listening, this is a strategy of last resort, one that is not scalable for the routine needs of moving massive amounts of information forward and making it readily accessible to users in the future.

These examples suffice to show that the relevant features of a digital artifact could include more than the fungible information contained in an electronic file. Although the Kaypro example may seem facetious, the many issues attendant to digital art and literature are not. The questions that the preservation of these digital objects bring forth are similar to those that collection development staff members ask about preserving all sorts of digital information that is dependent on specific software programs for its intrinsic value.

In many cases, the answer for libraries will be that they are, in fact, primarily concerned with collecting, preserving, and providing access to the fungible informational content of digital objects. In that case, the "preservation through handling" scheme is a likely winner; digital information that is frequently used by patrons stands a better chance of being migrated and refreshed, and therefore is more likely to continue to be available in future generations, than is little-used digital information. Indeed, migration may turn out to be a much more frequently recurring problem than is refreshing, because "today's optical media most likely will far outlast the capability of systems to retrieve and interpret the data stored on them" (Conway 2000). Regrettably, it is easier and cheaper to refresh than to migrate. If libraries have reason to be hopeful in this regard, it lies in open, nonproprietary standards such as JPEG for images, MPEG for video, and SGML, XML, and XSL for textual data. There are still important data types for which no such standards exist (GIS data, for example). However, the trend over the last 20 years—accelerated significantly in the last decade by the advent of the World Wide Web—has been in the direction of support for nonproprietary standards, even in proprietary software (Rosenthal and Reich 2000).[25]

[25] That this general truth should not be taken for granted by the library community was recently and significantly demonstrated by the release of Microsoft's Windows XP, with its Internet Explorer 6 (which does not have native support for Java) and its Windows Media Player (which does not have support for the MP3 music format).

Another promising strategy under development is called LOCKSS (Lots of Copies Keep Stuff Safe). The basic principle of LOCKSS is preservation through proliferation. An article in *The Economist* described LOCKSS as follows:

> . . . a network of PCs based at libraries around the world and designed to preserve access to scientific journals that are published on the web. The computers organise polls among themselves to find out whether files on their hard disks have been corrupted or altered, and replace them with intact copies if necessary. ("Here, There and Everywhere," June 24, 2000, 92)

Curiously, this strategy flies in the face of the new logic of preservation and access in the digital world: that libraries should not build redundant collections in a networked environment, but should develop cooperative collection development and preservation strategies. Duplication of effort is seen as not cost-effective. In the digital realm, cold storage does not work. Born-digital artifacts will not benefit from living undisturbed in dead storage for years. On the contrary, digital artifacts seem to require preservation through handling. Such active intervention would seem to obviate redundancy. Nevertheless, there is plenty of room for experiments in this new environment, and if the future proves to be anything like the past, then many of our initial approaches to ensuring longevity will be stood on their heads.

With print originals, the "preservation through proliferation" strategy of duplicative collections assumes that one copy is as good as another, that is, that the value of the artifact is fungible and that it is the information, not the artifact, one wants to preserve. One copy cannot serve the access needs of a large research community, and the only way for two people to have access to a book is for each to have his or her own copy. The copies cannot be shared in real time: If I give you mine, I don't have one. With digital files, the opposite is true: If I give you the file for an e-book, you have it and I have it, too. But a million perfect copies will still stand mute if we no longer understand how to read them.

3.4.6 Copyright: A Barrier to Preservation?

Libraries and archives are allowed to use copying technologies for preservation purposes under an exemption of the copyright code. When the Digital Millennium Copyright Act first appeared, there was no exemption for digital materials. The copyright community feared that it would be impossible to control the distribution of copied files, and that even those created for educational and preservation purposes would be a threat to the intellectual property rights of their creators or publishers.

These concerns have been addressed in further legislation that allows preservation copying, and technologies to encode data are adding desired layers of protection from infringement for rights owners. The real threat to preservation in the digital realm is that copyright law is being finessed by licensing agreements. Libraries do

not purchase electronic databases and do not have ownership of materials as they do of analog materials. Libraries do not preserve materials that they do not own. At the same time, however, publishers are not in the preservation business.

The library and publishing communities are making efforts to come to terms with the crisis that this situation may create (DLF 2001c). This is a matter that deeply affects scholars, as creators of intellectual property and as researchers and teachers wishing to gain access to the published record. It is one more area in which scholars must become familiar with what is at stake in negotiations between publishers and librarians and be sure that their interests are considered and that their responsibilities—as writers as well as teachers—are clear.

BEST COPY AVAILABLE

4. Case Studies

4.1 The Five-College Library Depository: *Combining Collections to Provide Better Preservation and More Comprehensive Collections of Print Materials*

The members of the Five Colleges of Massachusetts—Amherst, Hampshire, Mount Holyoke, and Smith Colleges, and the University of Massachusetts at Amherst—are developing a depository library that will house collections from each of their holdings. They are also creating a cooperative collection-development strategy that will take full advantage of the depository library.[26]

In deciding to take part in this joint project, each of the libraries faced the following constraints:

- a critical shortage of space to house growing collections
- a reluctance of its governing body to build separate storage facilities
- the need to implement the most cost-effective strategies for meeting users' needs

The colleges' effort to collaborate on collection development and management, as well as storage, can serve as an exemplar of the strengths and weaknesses of collaborative approaches to problem solving.

The cooperative depository will be built at the Amherst Library Depository facility, a former military bunker that was designed to survive a nuclear attack and to serve as a center for military operations in the event of a war. The facility was decommissioned long ago, and Amherst College purchased the bunker from its second owner, the Federal Reserve. The college retrofitted the bunker for secondary storage and began using it in 1995.

[26] This case study is based on *A Collaborative Approach to Collection Storage: The Five-College Library Depository* (Bridegam 2001).

The Five Colleges, in the words of the director of one of their libraries, "having taken successful consortial approaches to ordering, cataloging, circulation, subscription database management, and materials delivery in earlier collaborative efforts, explored ways to extend their cooperation to the growing problem of finding additional space to store little-used books in their collections." When the libraries first talked about cooperating to solve their storage needs, they considered several options, from building a new facility to renting space from a local commercial storage facility or from the nearest library storage facility (in this case, Harvard University's facility, which is about 100 miles away). Careful cost estimates led the library directors to recommend developing something more radical and more efficient—a shared library depository at Amherst.

In 1999, the presidents of the Five Colleges approved a plan to operate a consortial library depository that would have a distinctive mission and governance, and they assigned responsibility to an existing consortium, Five Colleges, Inc., to run the depository. There would be a consortial ownership of materials sent to the depository (except for the holdings of the University of Massachusetts at Amherst Library; as noted below), and duplicates would be deaccessioned. Materials would be housed by size, and the stacks would not be open for browsing. The existing Five College online library catalogs would be updated to show the new locations of materials transferred to and retained by the depository. Materials deposited by the four colleges would become the property of the Five-College Library Depository. Because the University of Massachusetts, a public research institution, is required to retain ownership of its materials, these would be shelved separately at the depository.

What obstacles did the colleges face in launching this collaborative endeavor? The first was to overcome the reluctance of some librarians and faculty members to transfer any materials off site. This opposition was not unexpected; nonetheless, in many respects, the colleges' options were limited. Their collections were growing, but budgets were tight. None of the institutions' governing bodies was willing to authorize funds for capital expenditures to house their libraries on site. Thus, the real issue was how the librarians and faculty could work together to make the transition equitably and with minimal disruption. In the case of Amherst College, for example, deciding which materials to move to the depository required extensive collaboration between librarians and the faculty. When conflicts arose among faculty members—usually over materials of an interdisciplinary nature—the librarians negotiated a settlement.

Another obstacle was that of ownership. Among the first decisions was to move to storage the journals that are available online, such as those mounted on JSTOR. By moving sets off site, the Five Colleges would be able to create a complete run of any given journal. By relinquishing ownership of that journal to Five Colleges, Inc., they would be agreeing to share ownership of and responsibility for common assets. Doing this might be easier for liberal arts colleges than for large private research institutions, which feel a need to keep

ownership of library materials. However, the arrangement worked out by Five Colleges was flexible enough to accommodate the specific ownership needs of a public university.

A series of logistical concerns had to be addressed, from document delivery to service on weekends and holidays. There were staffing and funding issues as well. As the collaborative effort moved forward, the colleges found it easier to solve their problems by pooling resources. For some issues, such as staffing needs, the flexibility afforded by a large body would prove to be essential.

The consortium established a Collection Management Committee to examine an array of matters, such as how to decide on acquisitions, what to do if one library decided to cancel a subscription that it alone had, and what materials should receive priority for deposit. The committee conducted research in most academic disciplines on the five campuses to determine the access needs and preferences of their faculties. It developed strategies to anticipate the recall of collections in cases of changing teaching demands. Finally, it determined that the long-term goal of the depository library would be to develop "last-resort" collections of best copies, so that no user would be without recourse to a hard copy of a title, yet the libraries would not be burdened with excessive duplication. The Librarians' Council, which comprises the head librarian of each college and the Five Colleges, Inc., coordinator, makes decisions about the depository on the basis of the committee's recommendations.

What can this example of cooperative collecting and storing offer to other libraries? First, libraries of all sizes have an increasing need for off-site storage. Large libraries have been moving collections off site for two decades, and several have built facilities that serve their collections alone.[27] (Some of these institutions—Harvard University is one—are also willing to lease some of their underused storage space to other libraries.) Many facilities have been built to serve several institutions in a region. Nevertheless, only one facility, the Midwest Inter-Library Center (renamed the Center for Research Libraries), was founded to manage cooperative acquisitions and preservation programs as well as to serve cooperative storage needs. The Center for Research Libraries has evolved from a model intended to serve the collecting needs of Midwest libraries by focusing on rare and little-used materials to its present profile, in which it serves 221 university and research libraries in North America and provides them with access to heavily used foreign language materials, newspapers, and documentary series.

The advent of digital technology and document delivery by fax have made the environment for collective storage far different today than it was in the 1960s and 1970s. No longer do scholars need fear that materials moved off site will be lost for casual use. Turnaround time at the Amherst College Library Depository is usually less than

[27] A forerunner of the Five-College Library Depository, the Hampshire Interlibrary Center [HILC], was established in the 1950s. HILC was disbanded during the 1970s, by which time each of the contributing libraries had erected larger libraries and had reclaimed their collections.

BEST COPY AVAILABLE

24 hours, and an increasing number of patrons are satisfied with desktop delivery. As long as the concept of secondary storage has been accepted on campus and users are satisfied with the levels of service provided, library managers can focus on how to use the shared repository to better serve the needs of both users and collections.

Whereas regionally based depository libraries have been seen chiefly as a way to handle a space problem, the opportunity now is to achieve economies of scale for a number of collection-management tasks, from acquisitions to preservation. Willis Bridegam, director of the Amherst College Library, states that even for the Five Colleges, the chief incentive to cooperate was to solve the problem of space shortages.

> They understood the economies that might be realized through joint staffing of a shared off-site library storage center. They saw the potential advantage of being able to develop complete periodical backruns from fragmented sets of the five individual libraries. They supported the idea of choosing the best copy of a book or periodical volume of which there were duplicates for retention in a depository. They also thought that it would be efficient to establish one conservation service at the bunker for all the materials transferred there. Most of all, they were interested in relieving the shelving space pressures in their libraries, and they thought that a joint approach might be more likely to attract external and internal funding (Bridegam 2001, 17).

Reaching agreement about the shared storage site, in other words, was aided and abetted by the press of short-term needs. The participants' experience of cooperation in sharing collections and cataloging, which had taken place over decades, had built trust among the libraries. The Five College presidents encouraged collaboration and supported innovative problem solving. This project bears watching, and it should be well documented and assessed regularly by the libraries and their users.

4.2. *The Emperor Jones*: *When Preserving Means Restoring*

Preserving films of historical or artistic value often entails a physical and historical reconstruction of what the film was in its original state and judging that state on the basis of evidence from as many authentic source materials as possible. That means that the restoration of a film may depend on the preservation of both film and non-film source materials that contain information about the film in its so-called original manifestation. These sources would include documents that might reveal what the film looked and sounded like, e.g., negatives, positives, scripts, stills, publicity materials, contemporary reviews, production company records, and copyright deposits. It also means that film preservation can require good detective skills and well-informed judgments about the cultural forces that shaped

the film at the time of its creation and later, as well as technical expertise in obsolete media, editing skills, and access to specialized equipment and reliable sources of funding.

In 1999, the Library of Congress undertook a restoration of *The Emperor Jones* as part of its contribution to the Treasures of American Film Archives project, which was initiated by the National Film Preservation Foundation with funds from the Pew Charitable Trusts. The Library staff's first task was to determine what constituted the original film when it was released in 1933.[28] There were several versions available, each of them incomplete in one or another way. The version of the film best known to the public was released by the American Film Institute in 1969–1970. It had been assembled from two pre-World War II-censored 16-mm prints. Another version, derived from a heavily censored source in Canada, was distributed in videotape format by Janus Film.

Preservation work began with gathering all extant versions and researching nearly every aspect of the film's production and distribution history. In this case, it is a history riven with controversy. That the film exists in such bowdlerized form today stems not only from the fact that many considered its content offensive or objectionable but that the lead actor, Paul Robeson, became a persona non grata because of his outspoken left-wing politics. Cultural artifacts associated with such controversy are particularly difficult to preserve, yet it becomes especially important to preserve them precisely for those reasons.

The film is based on an expressionist-style one-act play by Eugene O'Neill, first produced in 1920 and revived four years later. The 1924 revival starred Paul Robeson. The play was successful, if controversial, among critics and theatergoers. The subject matter and language were racially charged, but the poetic vision and expressionist staging made for compelling theater.

A small independent company produced the film in 1933. The producers secured O'Neill's permission to make the film, but the playwright had no involvement in it. His only requirement was that Robeson again play the lead role. (Robeson's character was Brutus Jones, a Pullman porter who becomes the lord of a Caribbean Island and ultimately meets a violent death.) The film script, written by DuBose Heyward, had received O'Neill's approval.

The film debuted in September 1933 and had a controversial opening run. White and Black audiences generally acclaimed the film a success, while agreeing it was more of an "art film" than entertainment. The African-American press voiced objections, saying that the lead character conformed to negative stereotypes of African Americans and that the language was replete with racial epithets.

[28] The substance of this case study was relayed by Ken Weissman, head of the Motion Picture Conservation Center at the Library of Congress, who supervised the work described here. He generously made available records that documented the work and additional background materials, and offered invaluable advice on the general matter of film preservation and restoration. James Cozart and Jennifer Dennis, of the Library of Congress, and Annette Melville, of the National Film Preservation Foundation, provided additional expertise.

The film played for only two weeks before its distributor, United Artists, began cutting certain parts that the critics had found objectionable.

These cuts, which included excising the word "nigger," used repeatedly by Jones as well as some White characters; cutting depictions of Whites subordinated to or physically intimidated by Blacks; and removing some suggestive sexual scenes, were not the first attempts at censorship of the film. Even before its general release, the producers, under advisement from censors, cut a scene in which Jones murders a White guard while serving on a chain gang and dialogue in which Jones describes guard brutality in prison (MacQueen 1990). As a result of prerelease editing, about two minutes was cut from the film.

Other immediate post-release cuts included removal from the sound track of all instances of frankly offensive words and even the word "Light!" spoken when Jones orders his White flunky to ignite his cigarette (the word was blanked out on the sound track but the image is unaltered). There is a mix of cuts that could be inferred from rough or jumpy spots in extant versions, but little unambiguous evidence about which scenes or parts of dialogue have been cut and when.

Library of Congress staff inspected all known sources of the film, including elements held at the Library itself, the National Archives of Canada, and the Museum of Modern Art. Staff members contacted numerous people who might have reason to have an unknown print, including some associated with the original production. They also researched contemporary reviews to determine how long the film ran, and production company records to learn more about what had been filmed. They obtained a copy of the shooting, or continuity, script that had been sent to a censor in New York and now resided at the New York Public Library. Also examined was a video transfer of the copy held by Gosfilmofond in Moscow—an English-language version with German subtitles that the experts hoped might include missing scenes. It did not, however, have any new material in it. Staff was also able to locate the sound track on Vitaphone-style discs owned by a private collector who was willing to lend them.[29] Listed in order of footage used, the Library used (1) the original picture and track negatives from the Library of Congress in the Universal Collection; (2) the incomplete studio print, also in the Universal Collection; (3) an incomplete Canadian print owned by the Museum of Modern Art; (4) Vitaphone-style sound disks owned by the late David Goldenberg; (5) the National Archives of Canada archival negative; (6) the archival picture negative owned by Janus; and (7) a pre-World War II 16-mm print owned by Douris Films and housed in the Rohauer Collection.

By September 2001, the Library of Congress had produced a copy of *The Emperor Jones* that experts believe to be approximately 3

[29] According to Annette Melville, the existence of these sound discs at such a late date is unusual and may indicate that the film was shown in small towns that did not yet have new sound systems.

minutes shorter than the original 80-minute film. Some elements of the film were still missing or had badly deteriorated. To get the film in sync with the sound track and to restore damaged frames, lab technicians doubled some frames and held others in freeze frame. The sound track was a complete version of that of the original general release.

The costs of this kind of preservation and restoration work, if done at a commercial facility, would be $35,000 or more for the film portion alone. Sound work could cost another $10,000 or more. These figures do not include the costs of research and staff time. Like rare-book conservation, film restoration is a high-investment solution for an artifact of special value that is deemed to be endangered. It is not a treatment that all films need to or should receive. Like works on paper and recorded sound, moving image materials require degrees of discrimination among objects and the treatment—passive care or active intervention—that should be used. There is always a balancing act between the items perceived to have high intellectual or cultural value at a specific time, the fragility of the physical item, and the funds and other resources available to treat them. It is fair to say that, as part of a larger national initiative to fund film preservation and restoration, *The Emperor Jones* was one of several works that received treatment because the funds became available.

4.3 Preserving Oral Traditions

Although the technology for recording sound is scarcely 100 years old, it has radically transformed the ways in which societies communicate, create art, document their lives, and fill idle time and empty rooms. It has also spawned large and lucrative industries that record and disseminate music, spoken word, and ambient sounds, and has introduced undreamed-of complexities into a copyright regime that was designed to manage rights for textual materials. Nearly every aspect of sound preservation is affected by these transformations: what to preserve, how to preserve it, how to negotiate and manage complex rights issues, and how to make recorded sound accessible into the future. Even for materials that are not commercially produced and distributed and that are, in essence, unpublished, rights and access issues can have a powerful inhibiting force on preservation. There is no better example of the challenges of sound preservation than that presented by recorded folklore.

The technical problems of preserving sound that is recorded on fragile media such as wax cylinders or cassette tapes, dependent as they are upon playback equipment that is quickly made obsolete by emerging technologies, may appear at first blush to be the chief obstacle to preservation and access. This is far from true. Future access to folklore resources will be equally dependent on the two other legs of the three-legged stool of access: (1) how materials are organized and described or cataloged for easy retrieval; and (2) whether or not present and future users will have legal or ethical rights to look at, cite, or reproduce the original sources. Given the challenge of physi-

cal preservation, few libraries and archives are willing to invest in preserving collections that are uncataloged or not cleared for research or educational uses. Most recorded folklore materials are technically unpublished, just as are radio broadcasts, television soundtracks, interviews, and recordings of live dramatic and musical performances.[30] This means that the folklore materials are seldom well indexed (indexing being an investment in inventory control that commercial firms must make to manage distribution, but which many documenters will not do systematically). There are equally parlous problems with intellectual and moral rights: a clear audit trail of written informed consent seldom accompanies folk recordings, and the rights of both the documenters and the documented are often contested under those circumstances.

The American Folklore Society (AFS) and the American Folklife Center (AFC) at the Library of Congress recently convened a group to identify the barriers to preserving audio folklore collections and to develop strategies for overcoming those barriers. Academic specialists and curators were concerned about decaying physical collections, which they saw as primarily technical. "These were the familiar challenges of media degradation and format obsolescence that have eluded effective remediation for at least a generation. To capture living traditions on documentary media, field workers have been using a variety of media formats, none of which is favorable for long-term preservation and each of which has presented new problems of storage, longevity, and hardware dependencies" (CLIR 2001). These media include wax cylinders; wire recordings; aluminum, shellac, and vinyl discs; glass and acetate masters; digital audiotape (DAT) tape; and cassette tape. Audiotape lasts for 10 to 60 years. But how can one tell exactly how long it will last, and what can one do to slow the inevitable loss?

The AFS and AFC invited technologists, preservation experts, lawyers, and members of various folk communities to discuss the issues and provide guides to best practices. The answers that the technologists gave to questions about stable media were discouraging: From the technical point of view, they said, there will never be a stable recording medium. Sound recordings must be periodically copied onto newer and, one hopes, more stable, media. We must learn to manage and migrate collections regularly, and to live with impermanence.

What is the format favored for future preservation? Whatever it is, it cannot be analog. The experts acknowledged that analog sound recordings have the highest fidelity. The problem with analog is that the media on which the sounds are recorded degrade. The sounds must be frequently copied onto other media, and with every instance of copying, some information is permanently lost. For preservation, the highest-quality medium for reformatting has been quarter-inch analog magnetic tape on open 10-inch reels. This has been the stan-

[30] "It is in these recordings in which rights issues are the most complex and in need of study, and perhaps adaptation, as they relate to preservation" (Brylawski 2001, 2).

dard for nearly half a century. But these tapes do not last as long as do the underlying sources, such as vinyl long-playing (LP) records. Even worse, only two firms manufacture this analog tape, and its production is not considered a growth industry. The advantage of digital recording is that there is no loss of information. The disadvantage is the need for compression, which diminishes the sound quality. Even with the relatively high fidelity that can be achieved with digital media, compact discs (CDs) will need reformatting and migration in the not-too-distant future.

Reformatting fragile media will require a process of ongoing assessment and triage to identify materials in need of treatment. This process will be successful only to the extent that collections and items in them are well organized and well described. While the immediate future will see management and migration of recorded sound from fragile media onto more stable media, this will be done in the knowledge that even the more stable media are not permanent. Under these circumstances, the best strategy is to develop a risk-management plan that takes into account the end purposes of preservation. By assessing future access demands, we can make informed decisions not only about what the medium or carrier needs to achieve stabilization but also about future users and their needs.

Projecting the access needs of future users is a crucial part of any preservation strategy. One cannot afford to wait and see what will be deemed valuable 25 to 50 years hence. In preserving folk heritage collections, we can judge future needs only by the present generation of scholars, researchers, and the documented communities themselves who need access to the materials.

One of the key findings of the survey of collections undertaken by the AFS and AFC is that much of what has been recorded is poorly controlled, badly labeled, and lacking critical documentation about rights to use. This situation exists in part because of the changing mores of professions such as anthropology, ethnomusicology, and folklore, which, decades ago, did not understand the importance of securing written or recorded informed consent from the peoples whom they were recording. There is no consensus on how to remediate this problem with retrospective collections; however, it is clear that scholars and field workers of the present and future must document fully the conditions under which their subjects grant access.

As with other academic disciplines, those of folklore, anthropology, and ethnomusicology are looking to the time when they might find both primary and secondary sources online. One of the major barriers is that the fields of folklore and ethnomusicology do not have standardized vocabularies for many common terms. This hinders the development of cataloging and indexing schemes for inventory control, not to mention searching across collections. It is critical that a thesaurus of terms be agreed upon, and work is currently being done on this. Issues of rights, vexing enough to sort through in the controlled environment of the reading room, become even more complex in the digital realm. Many people are calling for open and virtually unrestricted access on the Web to indigenous music and

folklore. Others urge that the oral traditions of various communities be controlled fully by those communities, even if that means banning access for research purposes by those who are not members of these communities. This has become a particularly heated issue among some Native American nations. In the meantime, there is a need to assess what individuals and institutions have on their shelves, what condition the materials are in, and how quickly these recordings must be reformatted to ensure that they are preserved.

4.4 JSTOR: *Online Access and Digital Archiving*

JSTOR, a nonprofit journal-archiving project, has made available online to libraries the back files of 153 titles from 17 academic disciplines, a collection that, as of November 2001, exceeded 8 million pages. More than 1,100 libraries from 53 countries participate in this collaborative enterprise. Approximately 180 colleges and universities have had access to the database since early 1997. In the first six months of 2001, more than 2.7 million articles were printed from its database, more than 5.7 million searches were performed, and the database had more than 22 million user access sessions (JSTOR-NEWS 2001).

The original intent of JSTOR was to reduce storage costs for low-use back journals in the humanities and social sciences and to ease access to journal content. While there is some evidence that academic libraries that use JSTOR are now either removing duplicate copies from their shelves or moving copies from the central library stacks to the more cost-effective shelving of remote storage, the chief benefit of JSTOR to date has not been for preservation. It has been to increase access to retrospective secondary literature and to dramatically change how these materials are used. Lexicologists and reference compilers are mining JSTOR, as they have mined the Making of America database. Fred Shapiro, compiler of the forthcoming *Yale Dictionary of Quotations*, has been using JSTOR to track down attributions for such quotable saws as "There is no such thing as a free lunch," and to determine that many such sayings have been incorrectly attributed for years by authoritative sources such as *Bartlett's Familiar Quotations* (Hafner 2001). (It was not Milton Friedman who first said this in 1975, but Alvin Hanssen in a 1952 *Ethics* article.) Others have been using the JSTOR database to pinpoint more precisely instances of first use of certain words (*Science* 2001).

It appears that the electronic articles in JSTOR are being used more frequently than are those in paper form. JSTOR studied the use of back hard-copy and electronic articles to compare access. There were a total of 692 uses of 10 hard-copy journals at five test sites over the course of the three-month survey in 1996. A study of the use of the same journals in JSTOR at the same five sites for the last three months of 1999 yielded more than 7,700 article views. In addition, although there is presumably substantial overlap in articles viewed and those printed, 4,885 articles were printed, bringing the total of articles viewed and printed during the study to 12,581. When this

figure is compared with the 692 uses in the 1996 survey, it would seem that electronic access is greatly increasing use of the material. Interdisciplinary use of the journals has also risen—a trend that has not been documented in hard-copy use. Further evidence shows that older articles in certain disciplines do not lose their value; for example, in the field of economics, the average age of the 10 most frequently retrieved articles was 13 years; in mathematics, it was 32 years. This is beginning to raise questions about what constitutes research or pedagogical value in journal literature and about the relationship between citation frequency and that value. The articles that are viewed most frequently are not the same as those that are most frequently cited; in fact, frequently viewed articles may be rarely cited (Guthrie 2000).

JSTOR was designed to solve both preservation problems and access problems. Its goal is to become a repository in which subscribers will eventually develop so much faith that they will relinquish many of their old and unused journals. It is not intended to spur a thoughtless disposal of journals, but rather to shift the burden of retaining low-use materials to those institutions that have taken on the role of "libraries of last resort." JSTOR has reached agreement with the Center for Research Libraries (CRL) to become the first North American repository of copies of every issue of every journal in its database. JSTOR intends to create several such repositories of hard-copy journals over time, in the belief that distributed artifactual repositories best serve the preservation needs of the research community through redundancy. Such repositories would also provide access to originals on demand. Many of JSTOR's advocates believe that a few full runs of journals centrally located in a few repositories are more useful than are a plethora of incomplete runs widely dispersed. In its work with libraries, JSTOR has often had difficulty in assembling complete runs of journals. This reveals how challenging it has been even for those libraries that try to be comprehensive in their coverage of one discipline or another (JSTOR 2001).

A number of economic issues may affect the further development of JSTOR and of its now-developing art image counterpart, ArtSTOR. These issues go to the heart of the promise of networked resources: to reduce costs resulting from unnecessary redundancies or, looked at in another way, to avoid future costs that libraries would incur if they continued to develop collections individually. As libraries' print collections grow, managers are forced to make decisions about where to store them: in the centrally located library stacks on campus or in the more flexible, preservation-friendly environments of offsite, high-density storage. They face two alternatives: (1) keeping a full retrospective collection onsite but not in all subjects and sending a great number of subjects off-campus; and (2) creating a more equitable burden among disciplines by finding the lowest-use items within a classification and sending them offsite. JSTOR statistics reveal that in some fields, the distinction in value or use between older and more recent literature is not meaningful, and using dates

of publication to predict demand may unintentionally create problems.

JSTOR has attempted to find out what it costs libraries to store back files of journals. One source has estimated that it costs a library an average of $175,000, based on what it would cost to build that storage today (Bowen 2001). Although this figure is derived from a methodology widely used in library science, it can be misleading because it is really about avoiding building in the future, not savings on current expectations. The costs of doing business in a research library are hard to calculate, since benefits are difficult to quantify. Nonetheless, the cost of housing journals onsite versus offsite can be clearly assessed. A survey that JSTOR conducted in 1999 revealed that 20 percent of respondents already had some journals in remote storage and that 24 percent of respondents indicated that they had plans to move more items offsite (Bowen 2001). Some libraries, especially small college libraries that are committed to serving the needs of teaching first and research second, have never had a policy of keeping journals forever. They are finding that they can gain access to older literature through JSTOR that they would have otherwise had to obtain through interlibrary loan. For libraries of last resort, as well as those attached to research institutions, the trade-offs between storage (onsite or offsite) and deaccessioning may differ and be harder to quantify. Few, if any, ARL libraries are divesting themselves of old hard copies, even the third of fourth copy, of a journal title they hold. For them, the key advantage of participation in JSTOR to date is improved delivery of resources to patrons. These libraries are not looking to eliminate storage costs at present. With JSTOR, however, they can send the second, third, or fourth copy to offsite storage with less concern about compromising ease of access. They are also able to provide access to journals with a convenience that would not otherwise have been possible. JSTOR has done something that no single library would have done on its own behalf: it has been willing to run the copyright gauntlet with publishers and arrive at an access policy that suits the needs of both publishers and researchers. ArtSTOR is aiming to accomplish the same thing with its database of art historical images.

4.5 The Rossetti Archive: *Collecting and Preserving the Born-digital Scholarly Publication*

The Complete Writings and Pictures of Dante Gabriel Rossetti: A Hypermedia Research Archive (the Rossetti Archive) is a comprehensive electronic edition produced and updated continually since 1993 by Jerome McGann and more than 30 others working under his direction.[31] The current version, published by the University of Virginia's Institute for Advanced Technology in the Humanities (IATH) using Enigma's Dynaweb software, is the first of four projected in-

[31] John Unsworth, director of the Institute for Advanced Technology in the Humanities at the University of Virginia, contributed this case study.

stallments. It includes 10,388 SGML and JPEG files, presenting material that centers on the 1870 volume of Rossetti's *Poems* and outlining the structure that the completed archive will require. This material is marked up in a Document Type Definition (DTD) developed for the project at IATH—the Rossetti Archive Master Document Type Definition (RAM DTD). In addition, there are about 5,000 (offline) TIFF images, from which the JPEGs are derived; some HTML pages with introductory, summary, and navigational materials; and perhaps two dozen style sheets. The publication also includes 18 essays about the archive, by McGann and others, marked up in HTML and available from the "Resources" area of the archive. The completed Rossetti Archive is likely to contain 25,000 files and to take another 10 to 12 years (and another 30 or 40 people) to finish. The University of Virginia, private foundations, and corporations have already invested hundreds of thousands of dollars in developing this resource; perhaps as much as a million dollars will be invested by the time the project ends.

The Rossetti Archive is a valuable scholarly publication, not only in terms of the effort and money invested but also in terms of the role it has played in the migration of humanities scholarship online, in the pioneering of electronic scholarly editions, and in the history of humanities computing. For these reasons, it is worth collecting and preserving in a research library. To date, however, digital library efforts have focused on library-based production of library-owned digital primary resources. Libraries have not yet had to deal with second-generation digital library problems, where the focus is on scholarly analysis, reprocessing, and creation of digital primary resources.

In January 2000, IATH and The University of Virginia Library's Digital Library Research and Development group (DLR&D) began a three-year project, "Supporting Digital Scholarship" (SDS), to investigate the problem that the Rossetti Archive and other originally digital scholarly publications pose to research libraries. Funded by The Andrew W. Mellon Foundation and codirected by IATH Director John Unsworth and DLR&D Director Thornton Staples, SDS aims to address second-generation digital library problems. This project is examining three digital library problems:

- scholarly use of digital primary resources
- library adoption of originally digital scholarly research
- co-creation of digital resources by scholars, publishers, and libraries

Approaching these problems requires developing technical methods and institutional policies for collecting originally digital scholarly publications. Accordingly, SDS has formed two working committees—one on technical issues and one on policy issues. The technical committee is responsible for production and implementation of the software, standards, and systems that this project requires. The goal of the technical committee is to build the systems that show what can and cannot be done, at a technical level, to support digital scholarship. The policy committee is charged with considering and

proposing policies governing long-term preservation and access for digital materials in the library and policies covering the integration, dissemination, and reuse of those materials. The goal of the policy committee is to produce guidelines for collecting digital scholarship that outline what libraries can and cannot promise to do with these materials, depending on what form they take, what standards they do or do not adhere to, what functionality they have, and how they achieve it.

Technical work in this project takes place within a digital library architecture called FEDORA (Flexible Extensible Digital Object Repository Architecture). FEDORA originated at Cornell University in research done by Carl Lagoze and others; the Virginia implementation is FEDORA's largest testbed to date and its first real-world installation. The Andrew W. Mellon Foundation has funded an extension of this research that will involve beta-test installations of FEDORA in a half-dozen research libraries in the United States and the United Kingdom.[32] In Virginia's implementation of FEDORA, objects within a digital library or repository consist of a basis (e.g., a JPEG image in a simple object, or a machine-readable text with page images, in a complex object), plus three metadata packages (administrative, technical, and descriptive). Finally, objects can be associated with one or more "disseminators"—data structures that pair a particular set of behaviors, or "signatures," with methods to produce that behavior (in the current IATH implementation, servlets).

In October 2001, the Rossetti Archive was first collected into FEDORA as a set of XML documents with XSL style sheets that mimic the functionality of the Dynaweb publication. To accomplish this, the SGML DTD had to be modified so that it was capable of validating SGML and XML. James Clark's SX was used to generate the XML, and mark-up in the documents themselves had to be adjusted to disambiguate some forms of references to other documents on which navigation and selection would depend. Entity declarations, held separately in a catalog file in the original SGML version, had to be distributed into the files that contained the relevant entity references. This task required the efforts of several staff members, working part-time, over several months—far beyond the effort that a library would be willing to devote to collect a single publication under normal circumstances, but far less time than has gone into creating the Rossetti Archive. The point of this experiment is not to demonstrate a cost-effective collection strategy but to develop an understanding of what characteristics an originally digital publication should have in order to be collectable at reasonable cost. Any such publication will have a better chance of being converted to a collectable form if it is highly structured, even if its form is idiosyncratic.

The publication's basic content—its textual information, image data and, perhaps most important from an editorial point of view, the profusion of relationships among the texts and images that make up the archive—is part of a library collection. It has been collected in

[32] Virginia's FEDORA implementation is outlined in Staples and Wayland 2000.

a way that should make it possible to migrate the data forward as mark-up standards, delivery mechanisms, browsers, and other elements of the digital library environment continue to develop.

Work remains to be done. The objects in the Rossetti Archive require administrative, technical, and descriptive metadata, and it is unclear how much of these metadata can be shared, how much can be automatically harvested from the existing data, and how much will have to be created by hand. There are some unsolved problems, too. Searching is one of the most difficult technical problems, because so much basic navigation (in the Rossetti Archive, but also in many SGML- or XML-based document structures) is predicated on queries, and because, until recently, there has been no standards-based way of expressing such queries. That problem may be solved for the Rossetti Archive by Tamino, a commercial XML database product that implements the full XPath standard and promises to implement XQuery, as soon as that standard is approved. If this works as advertised, there will be no part of the Rossetti Archive that cannot be expressed in a completely software- and hardware-independent way. In principle, therefore, all of its informational content or functionality will be fungible across future changes in technology.

That optimistic assessment should not obscure the difficulties that might attend the collection of a different sort of originally digital scholarly publication. For example, if the relationships in the Rossetti Archive had been embodied in a relational database, rather than in an XML structure, there would be no standards-based way to express them. Consequently, this important aspect of the publication could be expressed only in a software-dependent way. This would make it much more difficult, perhaps impossible, for a collecting library to make commitment to maintain, migrate, preserve, and access the materials over time.

If scholars are to produce originally digital publications that are compatible with libraries' needs and that allow libraries to collect at reasonable costs, then best practices for authoring have to be understood, established, and supported in some kind of networked, institutionalized work space. Publishers should ultimately support this workspace, though it may need to be designed and developed, initially, by libraries themselves. Prototyping such a workspace will occupy the second half of the SDS project.

5. Summary And Recommendations

> I learn with great satisfaction that you are about committing to the press the valuable historical and State papers you have been so long collecting. Time and accident are committing daily havoc on the originals deposited in our public offices. The late war has done the work of centuries in this business. The last cannot be recovered, but let us save what remains; not by vaults and locks which fence them from the public eye and use in consigning them to the waste of time, but by such a multiplication of copies, as shall place them beyond the reach of accident (Jefferson [1791] 1984).

Thomas Jefferson, often credited with having foreseen the problems his compatriots would face and devising solutions to them, seems in this letter to be typically prophetic in his vision. Is he not describing "preservation through proliferation"—the key preservation strategy that libraries adopted in the nineteenth and twentieth centuries and that has resulted in the great abundance of resources in libraries? Is not "multiplication of copies," as Jefferson puts it, the way that films and music and photographs have spread through our culture, with demand for preservation following? Is not copying them, be it by photocopying or scanning, and placing them beyond the reach of accident a recipe for avoiding the "daily havoc on the originals"? And, as we look ahead, will these remain the primary strategies for ensuring future access to present-day originals, or should we anticipate major changes in the nature of research, teaching, and preservation-and-access technologies that will render these strategies obsolete?

Core principles of artifactual value do not change with genre—map versus manuscript versus musical performance. The value of an artifact, however, is profoundly affected by the medium on which the information is fixed. The definition of artifact as a unique item of

historical importance worked as long as there was a fair degree of consensus regarding the nature of the artifact. New recording techniques for sound and image, as well as magnetic and electronic reproduction, have put enormous pressure on these heretofore-useful assumptions. This is true even in the case of traditional print/paper artifacts, as has been seen.

Artifacts are complex objects—materially, structurally, temporally, and perceptually. Cultural forces shape this complexity in an ongoing dialectic that is subject to continual revision and thus is never definitive. The status of any given artifact, like that of any cultural construction, is vulnerable for several reasons. In a museum or a library that has limited economic resources for core mission work, it must compete to maintain a claim on institutional resources. It must also compete for intellectual resources—for the attention of the public, of scholars, and of others who, by their attention or lack thereof, valorize its status. This intellectual competition is every bit as parlous as is the economic one. Each generation must engage the issue on its own terms and do so actively if preservation is to be effective.

Today, thanks to the very technologies that highlight, if not create, the problems we are facing, we can coordinate efforts to document and preserve in ways hitherto impossible. We should be able to track more clearly what others are doing in identifying what must be saved and what need not. We should be able to coordinate retention of best surviving copies and thereby avoid duplicating preservation efforts. We should be able to know with certainty what "last, best" copies of an artifact exist and where they exist, and therefore to assign responsibility for stewardship of such items.

This work is traditionally seen as the purview of librarians and archivists; however, preservation strategies at the local and national levels have been and will continue to be dependent for their success on the engagement of all members of the research community, including scholars and academic officers. Scholars may not see preservation of research collections as their responsibility, but until they do, there is a risk that many valuable research sources will not be preserved. Faculty can work with local librarians to ensure that scholars' research needs are clearly articulated and are taken into consideration when budgets are planned. Faculty can be influential in persuading presidents and provosts to devote appropriate levels of resources to locally held research collections. Above all, they can see that part of their role as scholars is to be stewards of collections as shared resources across the country. This means engaging at the national as well as local level to ensure that scarce but valuable resources are being preserved at some locations. They can engage their national roles as stewards through their scholarly societies, which act as publishers and guardians of the scholarly record. They can, through national organizations, articulate the value of the intellectual heritage that is at risk and lobby for increased funding to support preservation.

5.1 Principles of Good Stewardship

Good stewardship begins with scholars and librarians taking responsibility for the preservation of artifacts in the following ways:

- Scholars can work with librarians in identifying materials to preserve, focusing first on those resources that are most endangered.
- Scholars can identify and define categories of resources and work with librarians to locate the finest and best-preserved specimens.
- Scholars and librarians can collaborate to develop realistic policies for retention and disposition.
- Librarians can work together to develop responsibilities for collections of last resort, including aggregating collections into regional repositories.
- Librarians can work together to refine existing bibliographic and collection records to make it possible for researchers to locate these preserved artifacts.
- Scholars and librarians can develop and nurture local, regional, and national collaborations that are sustainable and that balance the needs of present and future users.

5.2 Best Practices for Preservation of the Artifact

At the technical level, good stewardship rests on the following principles and practices:

- preventive care
- global treatment of collections, when available, for cost-effectiveness
- use of item-level treatment for items of special value and vulnerability
- special handling and security protocols
- surrogacy to prevent damage and extend access
- optimal storage facilities
- networks of repositories and remote storage facilities
- careful documentation of problems
- constant assessment of progress in key programs and sharing of data about them
- documentation of preservation treatments (e.g., deacidification, reformatting) on bibliographical records

5.3 Strategies for Specific Formats

As part of the network of libraries responsible for preserving and making accessible artifactual collections, individual libraries can play a critical role in a national preservation strategy by doing the following:

Print (pp. 19-30)

- identifying which genres and what types of printed materials need to be preserved
- using the most appropriate and cost-effective means (e.g., deacidification, reformatting, rebinding, repair) to ensure long-term access

- giving priority to the preservation of materials uniquely held at local institutions or generally rare

Audiovisual (pp. 30-41)

- assessing collection needs to set priorities for reformatting and other treatments
- stabilizing materials on paper (e.g., engravings, posters, and photographs on acid paper)
- reformatting materials onto stable media before information is lost
- saving original source materials (e.g., negatives)
- documenting preservation treatments on common databases and coordinating preservation treatments to avoid duplication

Digital (pp. 41-54)

- developing sound digitization selection criteria
- ensuring preservation treatment of underlying source materials
- creating registries of digitized collections
- defining what types of digital objects (e.g., electronic journals, Web sites) should be preserved and by whom
- lowering barriers to fair use of digital information
- creating digital resources and retrieval systems in nonproprietary systems
- developing best practices for migration
- exploring the possibilities of emulation

5.4 Recommendations

At the local and national levels, there are several actions that members of the research community can take to ensure greater access to original research materials in the future and to use available resources most effectively. Such actions include the following:

- Advocate for the development of regional repositories of artifactual collections that reduce duplication of effort, create economies of scale, and ensure that the greatest number of unique or scarce priority items are preserved and made accessible to researchers. Such repositories might be organized along chronological lines, with institutions specializing in certain periods; along disciplinary or linguistic lines; or along geographical (i.e., physical location) lines for consortial use.
- Advocate for the creation of a repository of record for American imprints that operates at the national level to ensure that at least one copy of materials deposited for copyright will endure.
- Create standardized descriptive practices that make information about resources readily accessible through searchable databases. This is especially critical for academic fields that lack best practices and controlled vocabularies in documentation and retrieval.
- Raise awareness of the importance of preservation with the expectation that foundations, federal agencies, and others will increase their financial commitment to it.

- Increase the funds devoted to preservation by gaining a larger commitment from host institutions, raising endowment funds, and securing grants from government agencies and foundations.
- Invest resources in the development of more effective and economical preservation processes.

5.5 Areas for Further Research

The research agenda that has emerged from this investigation into the role of the artifact in library collections is considerable for each of the three types of recording media discussed: print, analog audiovisual, and digital. The most pressing areas, largely nontechnical, include determining what materials are held in libraries and archives, identifying how researchers use source materials, and encouraging the use of primary sources and artifactual collections in research and teaching. In particular, there is a need to

- develop and apply methodologies to track the growing online use of large digitized collections
- gather and study data on the behavior of researchers online so that the tool kit developed for researchers' use will meet their needs
- gather data on the state of artifacts in nonacademic libraries and repositories
- research and develop curricular needs for the use of original sources
- increase media longevity studies and extend them to all new media, including digital

The task force also recommends that a similar investigation be undertaken into the state of preservation among rare and special collections. While excluded from consideration in this study, these collections are as vulnerable to unintended loss and destruction as are the artifactual and digital collections that have been the subject of this report.

6. References

Web site addresses were valid as of November 1, 2001

Association of Research Libraries. 2000. Holdings of University Research Libraries in U. S. and Canada. *The Chronicle of Higher Education* May 19: A23.

Association of Research Libraries. 1999a. Association of Research Libraries Data Tables 1998–1999. Available at http://www.arl.org/stats/arlstat/mrstat.html.

Association of Research Libraries. 1999b. *Association of Research Libraries Newsletter* 206 (October). Special Issue on Global Resources.

Ayers, Edward L., et al. 2001. The Valley of the Shadow. Available at http://www.iath.virginia.edu/vshadow2/.

Baker, Nicholson. 2000. Deadline: A Desperate Plea to Stop the Trashing of America's Historic Newspapers. *The New Yorker* July 24: 42-61.

Baker, Nicholson. 2001. *Double Fold: Libraries and the Assault on Paper.* New York: Random House.

Bearman, David. 1999. Reality and Chimeras in the Preservation of Electronic Records. *D-Lib Magazine* 5 (4). Available at http://www.dlib.org/dlib/april99/bearman/04bearman.html.

Bearman, David, and Jennifer Trant. 1998. Authenticity of Digital Resources: Towards a Statement of Requirements in the Research Process. *D-Lib Magazine* (June). Available at http://www.dlib.org/dlib/june98/06bearman.html.

Big 12. 1998. Scholarly Communication and the Need for Collective Action: A Statement by the Chief Academic Officers of the Big 12. Press release, August 7. Available at http://www.big12plus.org/pressreleases/scholar.htm.

Bouché, Nicole. 1999. *Digitization for Scholarly Use: The Boswell Papers at the Beinecke Rare Book and Manuscript Library.* Washington, D.C.: Council on Library and Information Resources. Available at http://www.clir.org/pubs/reports/pub81-bouche/pub81text.html.

Bowen, William G. 2001. At a Slight Angle to the Universe: The University in a Digitized, Commercialized Age. The Romanes Lecture for 2000, delivered at the University of Oxford, October 17, 2000. *ARL Newsletter* 216 (June): 1–15.

Branin, Joseph, Frances Groen, and Suzanne Thorin. 2000. The Changing Nature of Collection Management in Research Libraries. Available at http://www.arl.org/collect/changing.htm.

Bridegam, Willis E. 2001. *A Collaborative Approach to Collection Storage: The Five-College Depository Library.* Washington, D.C.: Council on Library and Information Resources.

Brockman, William S., et al. 2001 forthcoming. *Scholarly Work in the Humanities and the Evolving Information Environment.* Washington, D.C.: Council on Library and Information Resources.

Brylawski, Samuel. 2001. Preservation of Digitally Recorded Sound. Unpublished report for the Library of Congress.

Chapman, Stephen, Paul Conway, and Anne Kenney. 1999. *Digital Imaging and Preservation Microfilm: The Future of the Hybrid Approach for the Preservation of Brittle Books.* Unpublished report. Washington, D.C.: Council on Library and Information Resources. Available at http://www.clir.org/pubs/archives/hybridintro.html#full.

Commission on Preservation and Access. 1997. Canadian Research on Permanent Paper Answers Old Questions and Raises New Concerns. *The Commission on Preservation and Access Newsletter* 104 (Nov-Dec): 2.

Conway, Paul. 2000. Overview: Rationale for Digitization and Preservation, in *Handbook for Digital Projects: A Management Tool for Preservation and Access*, edited by Maxine K. Sitts. Andover, Mass.: Northeast Document Conservation Center.

Conway, Paul. 1996. *Preservation in the Digital World.* Washington, D.C.: Commission on Preservation and Access. Available at http://www.clir.org/pubs/abstract/pub62.html.

Council on Library and Information Resources. 2001. *Folk Heritage Collections in Crisis*. Washington, D.C.: Council on Library and Information Resources. Available at http://www.clir.org/pubs/reports/pub96/contents.html.

Council on Library and Information Resources. 2000. *Authenticity in a Digital Environment*. Washington, D.C.: Council on Library and Information Resources. Available at http://www.clir.org/pubs/reports/pub92/contents.html.

Cox, Richard J. 2000. The Great Newspaper Caper: Backlash in the Digital Age. *First Monday: Peer-Reviewed Journal on the Internet* 5(12). Available at http://firstmonday.org/issues/issue 5_12/cox/index.html.

Digital Library Federation. 2001a. Digital Library Standards and Practices: Benchmarking Digital Reproductions of Printed Books and Serials. Available at http://www.diglib.org/standards.htm.

Digital Library Federation. 2001b. Registry of Digital Reproductions of Paper-based Books and Serials. Available at http://www.diglib.org/collections/reg/reg.htm.

Digital Library Federation. 2001c. Preservation of Electronic Scholarly Journals. Available at http://www.diglib.org/preserve/presjour.htm.

Foucault, Michel. 1997. What Is an Author? Translated by Donald F. Bouchard and Sherry Simons. In *Language, Counter-Memory, Practice*, edited by Donald F. Bouchard. Ithaca, New York: Cornell University Press.

George, Gerald. 1995. *Difficult Choices: How Can Scholars Help Save Endangered Research Resources? A Report to the Commission on Preservation and Access*. Washington, D.C.: Commission on Preservation and Access.

Gertz, Janet. 1999. Is Digital Conversion Really Part of Preservation? *Abbey Newsletter* 23(2): 22.

Granger, Stewart. 2000. Emulation as a Digital Preservation Strategy. *D-Lib Magazine* 6(10). Available at http://www.dlib.org/dlib/october00/granger/10granger.html.

Greenstein, Daniel I. 2001. The Registry's Return. *CLIR Issues* 19:5. Available at http://www.clir.org/pubs/issues/issues19.html#registry.

Guthrie, Kevin. 2000. Revitalizing Older Published Literature: Preliminary Lessons from the Use of JSTOR. Paper presented at the Economics and Usage of Digital Library Collections Conference, University of Michigan, Ann Arbor, Michigan, March 23, 2000.

Hafner, Katie. 2001. A New Way of Verifying Old and Familiar Sayings. *The New York Times* (Feb. 1): E8.

Hayes, Robert M. 1985. Analysis of the Magnitude, Costs, and Benefits of the Preservation of Research Library Books. Working paper prepared for the Council on Library Resources.

Jana, Reena. 2001. The Guggenheim's New Legacy System. Digital Art Preservation. *Wired* (October): 64.

Jefferson, Thomas. [1791] 1984. Thomas Jefferson to Ebenezer Hazard, Philadelphia, February 18, 1791. In *Thomas Jefferson: Writings: Autobiography, Notes on the State of Virginia, Public and Private Papers, Addresses, Letters*, edited by Merrill D. Peterson. New York: Library of America.

Jewell, Timothy D. 2001. *Selection and Presentation of Commercially Available Electronic Resources: Issues and Practices*. Washington, D.C.: Council on Library and Information Resources.

JSTORNEWS. 2001. Available at http://www.jstor.org/news/2001.06/june2001.pdf.

JSTOR. 2001. Available at http://www.jstor.org/about/issues/index.html.

Kantor, Paul B. 1986. *Costs of Preservation Microfilming at Research Libraries: A Study of Four Institutions*. Washington, D.C.: Council on Library Resources.

Kenney, Anne R. 1996. The Disposition of Archival Source Documents After Digital Imaging. *Janus* 2: 78-86.

Library of Congress. 2001. How Does the Library Acquire its Holdings? Frequently Asked Questions. Available at http://lcweb.loc.gov/faq/25faq.html#5.

Library of Congress. 1997. *Television and Video Preservation 1997: A Study of the Current State of American Television and Video Preservation*. Vol. 1. Report of the Librarian of Congress. Written by William T. Murphy. Washington, D.C.: Library of Congress. Available at http://lcweb.loc.gov/film/tvstudy.html.

Library of Congress. 1994. Redefining Film Preservation A National Plan. Recommendations of the Librarian of Congress in Consultation with the National Film Preservation Board. Coordinated by Annette Melville and Scott Simmon. Washington, D.C.: National Film Preservation Board of the Library of Congress. Available at http://lcweb.loc.gov/film.

Lyman, Peter, and Hal R. Varian. 2000. How Much Information? Available at http://www.sims.berkeley.edu/how-much-info/.

Lyman, Peter, and Brewster Kahle. 1998. Archiving Digital Cultural Artifacts: Organizing an Agenda for Action. *D-Lib Magazine*. Available at http://www.dlib.org/dlib/july98/07lyman.html.

Lynch, Clifford. 1999. Canonicalization: A Fundamental Tool to Facilitate Preservation and Management of Digital Information. *D-Lib Magazine* 5(9). Available at http://www.dlib.org/dlib/september99/09lynch.html.

MacQueen, Scott. 1990. Rise and Fall of *The Emperor Jones*, 1933. *American Cinematographer* (February): 34-40.

Marley, S. Branson, Jr. 1975. Newspapers and the Library of Congress. *Quarterly Journal of the Library of Congress* 32 (July): 207-237.

Menne-Haritz, Angelika, and Nils Brübach. 1999. The Intrinsic Value of Archive and Library Material. *Digitale Texte der Archivschule Marburg* (5). Available at http://www.uni-marburg.de/archivschule/intrinsengl.html.

Modern Language Association. 1996. Statement of the Significance of Primary Records. Report of the MLA Ad Hoc Committee on the Future of the Print Record (1993–1995). *Profession* 95: 27-50.

Molyneux, Robert E. 1998. *The Gerould Statistics 1907/08–1961/62*. Available at http://fisher.lib.virginia.edu/gerould/.

National Archives and Records Service. 1982. Intrinsic Value in Archival Material. Staff Information Paper 21. Washington, D.C.: National Archives and Records Service and General Services Administration.

Panitch, Judith M. 2001. *Special Collections in ARL Libraries; Results of the 1998 Survey Sponsored by the ARL Research Collections Committee*. Washington, D.C.: Association of Research Libraries.

Reed-Scott, Jutta. 1999. *Preserving Research Collections: A Collaboration Between Libraries and Scholars*. Washington, D.C.: Association of Research Libraries.

RLG DigiNews. 2000. FAQ on Digital Collection Registries. Available at http://rlg.org/preserv/diginews/diginews4-6.html.

Rosenthal, David S. H., and Vicky Reich. 2000. *Permanent Web Publishing: LOCKSS (Lots of Copies Keep Stuff Safe)*. Available at http://lockss.stanford.edu/freenix2000/freenix2000.html.

Rothenberg, Jeff. 1999. *Avoiding Technological Quicksand: Finding a Viable Technical Foundation for Digital Preservation*. Washington, D.C.: Council on Library and Information Resources.

Rothenberg, Jeff. 1995. Ensuring the Longevity of Digital Documents. *Scientific American* 272(1): 42-7.

Safire, William. 2000. The Way We Live Now: 9-10-00: On Language: Old Guard. On Language. *The New York Times Magazine* (Sept. 10): 28.

Science. 2001. You Read It Here First. *Science* 291 (Jan. 5): 39.

Smith, Geoff. 1995. Access to Newspaper Collections and Content in a Time of Change. *IFLA Journal* 21: 282.

Staples, Thornton, and Ross Wayland. 2000. Virginia Dons FEDORA: A Prototype for a Digital Object Repository. *D-Lib Magazine* 6 (7/8). Available at http://www.dlib.org/dlib/july00/staples/07staples.html.

Storm, William D. 1998. *Unified Strategy for the Preservation of Audio and Video Materials*. Preservation Research and Testing Series No. 9806. Washington, D.C.: Library of Congress Preservation Directorate.

Tanselle, G. Thomas. 1998. *Literature and Artifacts*. Charlottesville Va.: The Bibliographical Society of the University of Virginia.

Task Force on Archiving of Digital Information. 1996. *Preserving Digital Information: Report of the Task Force on Archiving of Digital Information*. Washington, D.C., and Mountain View, Calif.: Commission on Preservation and Access and the Research Libraries Group, Inc. Available at http://www.rlg.org/ArchTF/.

U.S. Census Bureau. 2001. *Statistical Abstract of the United States 2000 (120th ed.)*. Austin, Tex.: Hoover's Business Press.

U.S. Department of Commerce. 1989. *Historical Statistics of the United States: Colonial Times to 1970*. White Plains, N.Y.: Kraus International Publishers.

Walls, David. 2000. Relative Cost for Three Preservation Modes. *Abbey Newsletter* 24(4): 75.

Web sites of organizations and projects noted in report

British Universities Film & Video Council. Available at http://www.bufvc.ac.uk.

Columbia International Affairs Online. Available at http://www.ciaonet.org/.

Kaypro Emulation. Available at http://www.yoy.org/kaypro/.

National Film Preservation Board. Available at http://lcweb.loc.gov/film.

National Film Preservation Foundation. Available at http://www.filmpreservation.org.

Rossetti Archive. The Complete Writings and Pictures of Dante Gabriel Rossetti: A Hypermedia Research Archive. Available at http://www.iath.virginia.edu/rossetti/.

7. Appendixes

APPENDIX I
Current Library Practices in Collection Development and Preservation

APPENDIX II
Charge to the Task Force

APPENDIX III
Selection for Preservation Criteria from RLG, Cornell, and Harvard

APPENDIX IV
Expenditure Trends in ARL Libraries, 1986-2000

APPENDIX V
Non-print Holdings in ARL University Libraries

APPENDIX VI
Comparative Costs for Book Treatments

APPENDIX VII
National Recording Preservation Act of 2000

APPENDIX I
Current Library Practices in Collection Development and Preservation

Libraries have traditionally built local collections to serve local patrons and have acquired material to meet the needs of their communities. For academic libraries, that has meant a focus on resources for teaching and research. Preserving the collections has been synonymous with preserving the information contained in them. Because libraries acquire physical objects for use and reuse (often under the "first-sale" doctrine of the United States Copyright Code), this has meant that they own physical property rights to the items on their shelves. (They rarely own the intellectual property rights.) They thus are not only preserving information but also preserving institutional property. While this may appear self-evident, it becomes an important economic consideration when libraries begin collecting materials too fragile to save for long, such as videotapes, or materials that have no physical form, such as electronic databases. It also matters greatly when libraries are urged to preserve resources that are of national importance but that receive little local use, as was the case in the brittle-books microfilming program of the 1980s and 1990s. When do national priorities trump local ones, especially when the funding is locally based? And how does funding affect preservation selection priorities?

Building Collections: Libraries and Archives

Research libraries, no matter how large, collect only a small portion of all the information created and disseminated at any given time. In the nineteenth and early twentieth centuries, most academic libraries were shaped directly by the research needs of their faculties. This resulted in some very rich veins of bibliographical ore, but it also generally produced holdings that were deep, but not broad, in coverage.

With the growth of all sectors of higher education after World War II, libraries and their collections boomed. Collecting became somewhat regularized, professionalized, and far more ambitious. Many libraries at large, research-oriented institutions made concerted efforts to collect in virtually all areas that their academic departments covered. Research collections were critical in recruiting graduate students as well as faculty, and universities scaled up their collecting activities to provide on-site access to their users.

Libraries selectively acquire currently produced items that will become the primary documents of tomorrow. To measure how the growth in book publishing has made selection more difficult, one has

only to note the substantial increase in the number of books published in the United States since 1880.

Historical Statistics of the United States: Colonial Times to 1970 (U.S. Department of Commerce 1989) and the *Statistical Abstract of the United States* (U.S. Census Bureau 2001) give the following annual output figures for the first year of each decade between 1880 and 1990:

Year	Books Published
1880	2,076
1890	4,559
1900	6,356
1910	13,470
1920	8,422
1930	10,027
1940	11,328
1950	11,022
1960	15,012
1970	36,071
1980	42,377
1990	46,738

These figures represent only titles produced in the United States. Research libraries, however, typically acquire titles from countries around the world. In 1992–1993, the United Kingdom published 86,573 new titles; by 1995–1996, the figure had increased by 24 percent, to 107,263. During the same period, Japan increased its book production 58 percent (from 35,496 to 56,221); Russia, 26 percent (from 28,716 to 36,237); and Italy, 20 percent (from 29,351 to 35,236) (Association of Research Libraries 1999b, 10). A recent analysis of library purchases shows that acquisition of foreign materials is holding steady, but the portion of imprints collected of the total published is declining (Association of Research Libraries 1999b, 10). The worldwide output is roughly 600,000, of which the average Association of Research Libraries (ARL) library purchases about 50,000.

Most materials in these collections, as well as those in all academic libraries, are text based. The collections are built along disciplinary lines and are created to serve research and teaching needs. It has fallen to specialized repositories to collect the plethora of information of research value that libraries do not routinely acquire. Such groups as historical societies and special libraries (e.g., the American Antiquarian Society, the Newberry Library) tend to collect in specific topics (e.g., American history, regional history, travelers' accounts) and formats (e.g., manuscripts, maps, photographs, musical scores), and to serve as critical supplements to the research libraries on campus. (There are quite a few specialized libraries on campuses as well.) In contrast to research libraries, many of the specialized repositories collect chiefly primary sources. They are distinguished from the book collections of libraries because they are usually unpublished (e.g., manuscripts) or rare (e.g., broadsides); furthermore, only

BEST COPY AVAILABLE

the secondary literature that is germane to those collecting strengths is acquired. Their collecting policies are shaped by research needs, to the exclusion of curricular ones.

Despite the recent appropriation of "archive" as a verb to mean "to store" or "to preserve," the traditional meaning of archives as a noun is narrower. Archives are institutions that collect records created in the course of an activity—for example, business records or government records. The collection strategies of archives are different from those of libraries. The National Archives and Records Administration, for example, is charged to collect, organize, and make available the records of the federal government. It collects information created by the government in the course of its business. This information includes not only paper and electronic records but also films, photographs, posters, and other visual and sound materials. Archives tend to merge the function of record management and true archiving for long-term access. They may acquire large caches of records that they will retain for stated periods of time for legal reasons. However, the average archives will accession for permanent retention only 1 to 5 percent of those records. During a process known as appraisal, they determine which files have historical value, and they keep only those materials.

Repositories that specialize in non-print materials are often called archives; however, film and sound archives do not have records in the technical sense because they are not documents created in the course of business. Their collections are scaled in an archival way: they are large and often described at the collection level. These special-format archives—some affiliated with universities, some with museums, some independent—are also a crucial part of the information landscape. They must play leading roles in any strategies the task force proposes to ensure the preservation of and access to artifactual collections.

Shared Access and Shared Collections

Until recently, the only way in which a library could make a publication available to its users was to own it. If demand were great, a library would have to own more than one copy. With the advent of photocopying and faxing, together with robust interlibrary loan and document-delivery systems, libraries can provide access to an item without purchasing it. Nevertheless, libraries still have large collections that replicate the collections of other libraries, in part because of the constraints of dealing with physically fixed information.

One of the chief advantages of digital technology is its ability to overleap constraints of time and place and to deliver information at any time to any computer that is connected to the Internet. Now that libraries no longer face the physical and temporal constraints that existed before the introduction of electronic networks, the models of access are changing significantly. Libraries understand that students and scholars prize instant access. Students seek out assigned readings a few hours before they are expected to have read them. Faculty

members track down bibliographical references under tight deadlines from publishers. As service organizations, libraries assign the physical locations of their collections on the basis of known patterns of use. The problems associated with keeping collections readily accessible have been highlighted in recent years, as more and more libraries are building offsite storage facilities for low-use materials. As economically compelling as offsite storage is, it succeeds only when faculty members are willing to recognize the time constraints built into the system (in many cases, a 24-hour wait between requesting the material and receiving it). But for resources such as journal articles, which can be delivered to the desktop at 3 a.m., how valuable to the scholar is the physical proximity of the hard copy? How much should a library spend to make both easily available? Faced with choosing between the two, how should the library decide which takes priority?

The new delivery technologies—from photocopy to interlibrary loan to digital access—are also having an effect on the substance of collection building. Do all libraries need to collect essentially the same materials, if they can be networked, either virtually or through interlibrary loan? Research library collections have been undergoing a certain homogenization, partly because of the spiraling costs of journals and partly because of the advent of buying consortia. Libraries have begun to admit that they can no longer aspire to collect comprehensively, both because of the escalating costs of journals and monographs and because of the increased quantities of materials. Between 1986 and 1996, the price of the average journal rose 147 percent, and that of a book by 63 percent. The number of interlibrary loans handled during the same period increased 102 percent. The access cost increased to an average of $30 per transaction. During the same period, the consumer price index increased by 41 percent. Citing these statistics, the Committee on Institutional Cooperation said in a 1998 press release, "We and our colleagues on the faculty must acknowledge that *research universities can no longer afford to build comprehensive collections.*" [Emphasis in original]. The committee called for new efforts of cooperative collecting (Big 12 1998).

The subject of cooperative collection development—the sharing of responsibility for important but low-use materials, many of them in foreign languages—is not new. When library collections first started growing rapidly after World War II, there was an attempt to share collecting responsibilities. When this attempt gave out, it was followed by a plan to build a national periodicals center in the 1970s. In the 1980s, the Research Libraries Group (RLG) developed a method, based on the Library of Congress classification scheme, that its members could use to inventory collections and determine areas of strength. The so-called Conspectus was an instrument that libraries used for a period of time to learn what others had collected, in part to reduce duplication of effort in certain specialized fields. Individual libraries declared their responsibility for collecting and preserving literature in areas they chose. During the heyday of filming brittle books, this inventory system was the basis for many decisions about

what each institution would film. The Conspectus fell into desuetude in the 1990s because, among other reasons, libraries were unable to keep acquisition information current.

These and other attempts to coordinate acquisitions at the national level have failed. There are, however, examples of successful regional models, one of which is the Triangle Research Libraries Network in North Carolina. Other examples include the Five-College Depository Library in Massachusetts (see Section 4) and the Center for Research Libraries.

User studies have repeatedly shown that many volumes in research libraries have extremely low use and or no use at all. A 1979 study of one medium-size research library showed that "any given book purchased had only slightly better than one chance in two of ever being borrowed. As books on the shelves aged and did not circulate, the chances of their ever circulating diminished to as low as one in fifty. Journal use, in general, was also discovered to be low" (Branin, Groen, and Thorin 2000). Given this, how could cooperative collection efforts fail?

Among the reasons generally cited for the failure of such efforts are the pressure for each library to serve local needs and its reluctance to cede responsibility in any one area to another library. Moreover, until recently, there were not enough commonly available bibliographical records to make such cooperation workable. Even today, several ARL libraries have not converted significant portions of their card catalogs into online records.

The tensions between local and national priorities for acquisition are mirrored in selecting for preservation. How should libraries manage the conflict between the present needs of researchers and the needs of future, or even physically distant, users? The challenge is to find ways to address the needs of local service today with the needs of the future and national priorities in mind. What should be the selection criteria?

Best Practices for Preservation Treatments

The fundamental principles of preservation apply to all formats, from print to digital, but the techniques and costs of the various treatments differ dramatically. The purpose of library preservation is to ensure the present and future use of information in whatever form it has been recorded. Library conservation is not like museum conservation, which aims to make an object fit for essentially passive use, such as exhibition. Library materials can be heavily used and must withstand the risk of misuse.

The great bane of libraries and archives is that there are no media on which information has been recorded that do not decay. As the carrier decays, the information goes with it. The optimal treatment is to preserve both the information and the carrier. In an academic library, for example, the ideal way to save a text has been to save the book or journal or whatever it has been printed to. Howev-

er, given the fragility of media such as wood-pulp paper, not to mention magnetic tape or onion-skin paper, this is not always possible.

A review of the diverse formats and media on which library collections are recorded led the task force to accept the following premises:

- All physical formats, from paper to magnetic tape, will decay over time.
- Physical handling compromises all physical formats.
- All copying from one physical medium to another, such as from book to photocopy or from LP to tape, results in some loss of information and will usually compromise the physical integrity of the original as well.

These are the facts that lead to a tension between preservation and access. In the world of physical objects, one is usually bought at the price of the other. The art of preservation is to minimize the risk of loss while continuing to keep collections usable for researchers. This entails the following:

- identifying the risks
- deciding what measures should be taken to mitigate those risks
- deciding which measures have priority

Little about preservation, other than its theory, is simple. Identifying risks can be tricky, and too often the risks become known only by the effects of decay, when it is too late to prevent loss. For example, we did not know much about the deleterious effects of light and humidity on wood-pulp paper until tests performed in the 1970s proved the connection. By that time, library collections were already full of books that had been printed on wood-pulp paper instead of the more durable, but also more expensive, rag paper. By the same token, we do not know the effects of car pollution on paper. Early test results indicate that car exhaust (as opposed to factory exhaust) is deleterious both to rag paper and wood-pulp paper (Commission on Preservation and Access 1997). If intervention is required to save an item or a collection of items, the timing of that action is also affected by factors that are hard to determine. As a rule of thumb, the newer the medium, the shorter its life span. Most new media, however, have not been tested well enough to make it possible to predict rates of loss; only when deterioration sets in do we begin to understand the life span. Who would have thought that the so-called safety film onto which nitrate film was transferred would itself deteriorate and fade in a few decades? And what does this mean for setting preservation priorities? Do we concentrate on stabilizing information on newer and more fragile media, such as reel-to-reel tape? Does that mean libraries should do so in lieu of deacidifying older books printed on acidic paper?

Other external factors enter into the calculation of risk: for example, funding and fashion, or supply and demand. Leaving aside money for the moment, demand for library materials that are robust enough to withstand use means that preservation departments must

first serve the needs of their local clientele. Much of the time and effort of preservation departments go to keeping the collections in circulation. Materials that are more frequently used receive more attention than do low-use items. Materials that are not frequently used, no matter how short the expected life span of the medium, receive short shrift. The Brittle Books Program, which started in the late 1980s, is the exception that proves the rule. When the larger research community become aware of the limited life span of many nineteenth- and twentieth-century imprints, the U.S. Congress allocated funds to the National Endowment for the Humanities to preserve the information in these resources that are, by definition, low-use. This has unfortunately not been true of feature and documentary films of the first half of the twentieth century, whose value for research was recognized too late for rescue. This fate has almost befallen comic books from the 1930s, and it will soon befall oral histories and ethnographic recordings on cassette tapes.

Deciding what measures to take begins with assessing the need, looking at the costs of treatment options, and looking for value for money. The most cost-effective means by far of ensuring the fitness of collections is to prevent damage. Preventive care is the heart and soul of most library preservation programs. The most cost-effective treatments are those that can be applied globally; one example is ensuring proper storage conditions. The decay that comes from the "inherent vice" of all physical media can be retarded through proper environmental conditions both in storage and in use. (The environmental needs [e.g., ideal temperature and humidity] of paper-based materials differ from those of film-based materials, and optimal storage requires separate facilities for each material.) Other global preservation treatments include the following:

- emergency-preparedness and disaster-recovery plans
- research and testing of materials (such as bar-code stickers that do not cause damage)
- development of standards and specifications for treatments
- education of staff and users about proper handling techniques

Collection-level treatments that act primarily to stabilize materials include the following:

- deacidification
- proper housing (e.g., using protective coverings, Mylar sleeves, and acid-neutral folders, or boxing a book that is fragile)
- binding loose materials

Item-level treatments are more expensive than are collection-level treatments because they require not only treatment of individual items but additional, often times skilled, labor to assess items and select them from within large holdings. Item-level treatments include the following:

- microfilming
- digitizing
- preservation photocopying

- creating photographic surrogates
- remastering sound recordings and moving images

If preservation fails, then conservation is the next step.* Item-level treatments include the following:

- book conservation, which may include repairing both the interior and exterior, removing the old boards, reattaching pages, reassembling and resewing the text block, making a new binding, and creating a customized box to hold all the source materials;
- paper conservation, which may include washing, mending, leaf-casting, backing, deacidifying, and reattaching fragments; and
- photo conservation, which may entail stabilizing edges, applying emulsion, creating copy prints for access, humidifying, drying, flattening, and mending and matting vintage prints.

The purpose of conservation is to return an item to use, even if that use is very limited. Preservation experts commonly say that preservation without access is not a good use of resources. A rare book, such as the *Bay Psalm Book*, may be so deteriorated that it cannot be opened without causing further damage. The goal of intervention in such a case is to render the book usable—that is, capable of being opened, even partially, under certain circumstances. In nearly all cases, the goal is not to restore an item to its "original state." That would generally entail recreating the book using modern materials with the same look and feel and technique. Such an approach compromises its authenticity. Age and deterioration are facts of life, and few preservation experts strive to return an object to its original form.

In part because funding for research and development efforts in library preservation has traditionally been so modest, treatments have been deployed in the past without proper testing, and these treatments themselves have turned out to be deleterious. Lamination is one example; the use of adhesive tape is another. Hence, contemporary experts honor two principles: first, do no harm; and second, do nothing irreversible. This means carefully documenting each treatment and testing materials to ensure that nothing damages original items.

In all formats, the guiding principle of selection for preservation treatment is to make an item fit for purpose. The act of choosing what gets treatment and deciding why is based on a constantly changing evaluation of the value, use, and condition of an item, and all of these factors are constrained by limited resources and inadequacies of technology. While the library and its personnel can be relied on to provide the expertise to assess condition, decide on treatments, track use, and so forth, the question of the changing valuation of the significance of artifacts for research and teaching is a matter that needs to be informed by scholars and teachers.

* "Preservation" is the generic term for all types of treatments, both preventive and corrective, that serve to stabilize items. "Conservation" usually refers to an item-level treatment that involves active intervention.

APPENDIX II
Charge to the Task Force

Rationale

Preservation has developed rapidly in the last two decades into a critical component in the life cycle of scholarly collection development and management, taking its place among selection, acquisition, cataloging, and service as one of the core functions of libraries and archives. Most research libraries allocate resources for the purchase of preservation services, employ professionally trained preservation specialists, and many have created whole divisions within an organization devoted to collections care. Moreover, preservation specialists now have a well-developed arsenal of tools and techniques available to them to apply to specific collection maintenance problems, ranging in scope and urgency from emergency response and environmental controls to mitigating wear and tear from use and slowing the self-destruction of various media.

Scholars and custodians, such as librarians and archivists, have mutual interests in the collection-building process. Although scholars are routinely and often intimately engaged with librarians in the acquisitions and use phases of collection building, they are usually exposed only to specific aspects of preservation activities (such as testifying to the value of a collection proposed for preservation microfilming) and rarely have the opportunity to view the preservation function as a whole.

However, recent vigorous debates in the scholarly community about the value of saving the book as an artifact in addition to preserving content from physically degrading books have raised important issues about scholars' knowledge of and participation in decision making about the disposition of original materials after preservation treatment. Through the creation of a task force, CLIR proposes to engage the scholarly community in a systematic review of its interests in this issue. The focus on preservation is part of a wider series of CLIR initiatives, launched in 1997 with the jointly-sponsored CLIR/ACLS task forces, intended to develop a forward-looking vision of collection building in an electronic information environment built on the valuable research collections made of print and audiovisual materials.

Charge

The Task Force is charged to articulate for scholars and librarians a general context or framework for formulating and/or evaluating in-

stitutional policies on the retention or disposal of published and archival or unpublished materials in the form the works were created.

The Task Force should answer the following questions in developing a general framework for informed decision making about disposition of originals:

- What factors make it useful and/or necessary to retain work in its original form? Under what circumstances are original materials required for research?
- When is it sufficient and appropriate to capture intellectual content through reformatting and not necessarily retain the original?
- What preservation options are available and what do they cost? From both custodial and scholarly perspectives, what are the advantages and disadvantages of these various options?

The Task Force should, to the extent possible, draw on actual experience and empirical evidence regarding scholarly uses of materials and different libraries' approaches to preservation. The analysis should look at the following preservation treatments: conservation, deacidification, off-site storage, preservation microfilming, and digitization.

Finally, the Task Force is specifically asked to consider the advisability and feasibility of creating one or more national repositories into which one copy of all materials published in the United States would be deposited for permanent retention.

Scope

The Task Force is asked to take its charge broadly, giving primary consideration to print formats, but also consider the burgeoning legacy of non-print and electronic research sources that demand increasingly urgent attention from preservation specialists. The Task Force should formulate its findings and recommendations in a way that takes account of these other, often more fragile, formats and suggests how the findings and recommendations might be extended to or tested in relation to them.

APPENDIX III
Selection for Preservation Criteria from RLG, Cornell, and Harvard

The Book as Object*
By the RLG Preservation Committee

For collection managers and curators reviewing materials, this list can serve as an aid when assessing library materials that might be rare or valuable. It explains why books become rare and deserve retention in their format.

Many items are important because of their format; often reasons are clear for maintaining those titles in their original state. In other cases, the reasons may not be so clear, but before withdrawing or converting to another format (due to deterioration, space-saving needs, superseded editions, or duplication) they should be reviewed. Hopefully, the considerations below provide an incentive to retain those items possessing valuable or important information in their physical format which might otherwise be lost.

The RLG Preservation Committee developed this list, with suggestions from two other RLG groups: the Collection Management and Development Program Committee, and the Archives, Manuscripts, and Special Collections Program Committee. Documents consulted were:

- The National Archives and Records Administration document *Intrinsic Value in Archival Material* (Staff Information Paper 21).
- *Transfer of Materials to Special Collections* of the Archives and Special Collections Task Force, Rare Book and Manuscript Section, Association of College and Research Libraries.
- An unpublished article, "The Preservation of Bibliographic Evidence," by Ellen McCrady.
- New York Public Library Technical Memorandum No. 40, *Permanent Retention of Materials in the General Collections in their Original Format*.
- The chapter, "Selection of Materials for Microfilming" in *Preservation Microfilming: A Guide for Librarians and Archivists* (Chicago: American Library Association, 1987).

This list is neither prescriptive nor presented in priority order. It does not represent RLG policy and is offered for informational, education, and selection aid only.

* Source: Nancy E. Elkington, ed. 1992. *RLG Preservation Microfilming Handbook*. Mountain View, Calif.: The Research Libraries Group, Inc. pp. 62-64. Reprinted with permission.

CONSIDERATIONS FOR RETAINING ITEMS IN ORIGINAL FORMAT

1. Evidential value—does the item possess or demonstrate:
   - The printing history of the item, such as registration pin marks, cancels, printing techniques, and typographic errors.
   - The binding history of the volume such as original sewing stations, binding structure, printed wastepapers used in the spine lining, and cover materials.
   - Marginalia, marks of ownership, and relevant ephemera laid or tipped in.

2. Aesthetic value—does the item have:
   - Bindings demonstrating:
     - unusual technique or artistry.
     - historical/developmental interest of structure or materials.
     - signed/designer bindings.
     - early publisher's bindings.
   - Other book decorations (e.g., gilding, gauffering, decorated endpapers, fore-edge paintings).
   - Illustrations not easily reproduced or meaningful only in the original color or original woodcuts, etchings, lithographs, etc.
   - Importance as an "artists' book" where the book is designed as an object.
   - Original photographs.
   - Maps of importance.
   - Pencil, ink, or watercolor sketches.

3. Importance in the printing history of significant titles—does the item possess any of these characteristics:
   - First appearance of the title.
   - Important bibliographic variants.
   - Important (or collected) fine press printings.
   - Technique important to the printing history.
   - Examples of early local imprints.

4. Age—determine if an item was:
   - Printed before [specific dates] in [specific countries] (e.g., all titles printed before 1850 in the U.S. or all books printed before 1801).
   - Printed during the incunabula period of any area (the first decades).
   - Printed during specific later periods, such as war years, in specific countries.

5. Scarcity—determine if an item was:
   - Rare in RLG member, NUC, and/or major European libraries.
   - Of fewer than 100 copies printed.

6. Association value of important, famous, locally collected figures or topics—does the item contain:
   - Notes in the margin, on endpapers, within the text.

- Bookplates and other ownership marks of such figures; other evidence of significant provenance.
- Important inscriptions and/or signatures.

7. Value assessed or sold at more than [specific cost].

8. Physical format or features of interest—does the item possess any of these characteristics:
   - Contains significant examples of various forms demonstrating technological development.
   - Exhibits unique or curious physical features (e.g., interesting watermarks, printing on vellum, wax seals).
   - Is an ephemeral material likely to be scarce, such as a letter-sheet, poster, songster, or broadside.
   - Contains some manuscript materials.
   - Is a miniature book (10 cm or less in height).
   - Is of questionable authenticity where the physical format may help verify it.
   - Is representative of styles, fads, mass printings currently rare.

9. Exhibit value—is the item:
   - Important to an historical event, a significant issue, or in illustrating the subject or creator.
   - Censored or banned.

Conservation Treatment: Library Materials to be Retained in the Collection in Original Format

Cornell University Library
Department of Preservation and Conservation
Reproduced with permission

The following criteria are designed to help identify library materials which should be preserved in their original format. In general, this material should not be replaced by reprographic means, reformatted, rebound, or repaired without careful consideration by the Conservation Liaison Specialist and the appropriate curators and bibliographers. The criteria should be applied to library materials regardless of shelf location, as it is recognized that circulating collections often contain materials which should be preserved in original format (see also RLG Preservation Manual, "Book as Object," pp. 62–64). Items identified for treatment from the general circulating collections to which the criteria apply should be brought to the attention of the Head of Rare Books (or appropriate curator) with a view to transfer to a sequestered collection.

1. Pre-1850 imprints and imprints issued after 1850 that can be considered rare or especially interesting because of time, place, and subject.

2. Materials having notable illustrations, maps, engravings etc, adding to the work's interest and/or value.

3. First editions of significant works.

4. Books with bindings of special interest because of type, period or binder.

5. Manuscripts and typescripts.

6. Association copies and materials having autographs of significance.

7. Material having added notes, annotations, marginalia etc. adding to the interest and or value of the work.

8. Materials in special closed collections.

9. Rare ephemera and other elusive items, such as songsters, sheet music, broadsides, almanacs, charts, original historic photographs, etc.

10. Notable standard reference works no longer in print such as dictionaries and encyclopedias.

In general, all books printed before 1850 and in original bindings will be restored rather than rebound, with all the features of the original bindings retained.

Library Preservation at Harvard: A Definition of Terms

Harvard University Library Preservation Center
January 2000 revision
Reproduced with permission

Conservation treatment, reformatting, commercial library binding, environmental control, disaster preparedness and response, and preservation education and training are the core initiatives that make up a comprehensive preservation program. Implementing such a program requires skilled staff, carefully selected and implemented technologies, and sound management practices.

CONSERVATION: Protecting and restoring the original object

The term conservation embraces activities that improve the condition of an object or protect it from damage. Paper treatments, for example, include everything from mending a small tear to washing, deacidifying, and other complex chemical and mechanical treatments. For books, treatments range from tipping in an errata sheet to full conservation rebinding. Conservation is also an appropriate strategy for preserving certain non-paper media. Conservation of motion picture film, for example, can include ultrasonic cleaning, splicing of breaks, and sprocket repair. Conservation activities also include such activities as pamphlet binding, boxing, enveloping, and other means of protective enclosure.

Collections conservation is the treatment and protective enclosure of materials that are valuable in the aggregate (e.g., the literature of a subject area or period, posters of a particular genre). An individual item is more important as it relates to other materials in the collection than it is as a disassociated object. Typically, 19th and 20th-century books and journals in circulating collections are the focus of a collections conservation program. While collections conservation techniques are often complex, items with similar problems can be grouped for batch treatment. Significant numbers of materials are conserved annually (tens of thousands in a large research library) using archival-quality materials and techniques. The goal is to preserve large collections of scholarly significance and therein to improve the library's overall service performance.

Special collections conservation is the treatment of materials that, while they may be important because of their relationship to larger collections, also have intrinsic value (and sometimes great monetary value) as isolated objects—much as do important pieces in museum collections. They may be valuable because they are rare, unique, were owned by an important person, are very beautiful or for other historical or aesthetic reasons. Conservation treatments involve extensive written and photographic documentation and the use of instruments such as powerful microscopes. The goal is to protect treasures having extraordinary local or worldwide significance.

REFORMATTING: Reproducing information at risk

Reformatting is the copying of library materials. Deteriorated objects—brittle books, for example—are copied onto more stable materials (silver halide microfilm, alkaline paper) to preserve information. Decaying nitrate negatives are copied onto stable film; motion picture film is copied to new film or to videotape to protect the original from the wear and tear of repeated use; photographs may be rephotographed using analog or digital means in order to protect originals from handling; sound recordings are copied to new media to ensure that they remain usable. Not all copying is motivated by preservation concerns. New versions are sometimes created to improve distributability and/or functionality. Scanned text that has been processed with Optical Character Recognition software is word-searchable and can be distributed widely over networks. Conversion of information to digital form can sometimes create rather than solve preservation problems.

COMMERCIAL LIBRARY BINDING: Managing the use of contractual services

Commercial library binding is the binding and rebinding of volumes in a largely automated commercial facility. In general, journal issues are bound together to collocate them and to prevent damage and loss; paperbacks are bound to make them more sturdy for library use; hardcover books are rebound because they are damaged and do not merit conservation treatment.

ENVIRONMENTAL CONTROL: Achieving hospitable storage conditions

Creating and maintaining storage conditions that promote the longevity of collections is the single most beneficial preservation strategy that a library can pursue. Controlling temperature, relative humidity, light, and air quality within appropriate ranges can radically slow the deterioration of paper, leather, cloth, plastic, and other materials ubiquitous in libraries. In general, cool, dry, low-light spaces free of gaseous and particulate pollutants are optimal. Ongoing monitoring of conditions is essential to maintaining preservation-quality conditions.

EMERGENCY PREPAREDNESS AND RESPONSE: Protecting collections from hazards

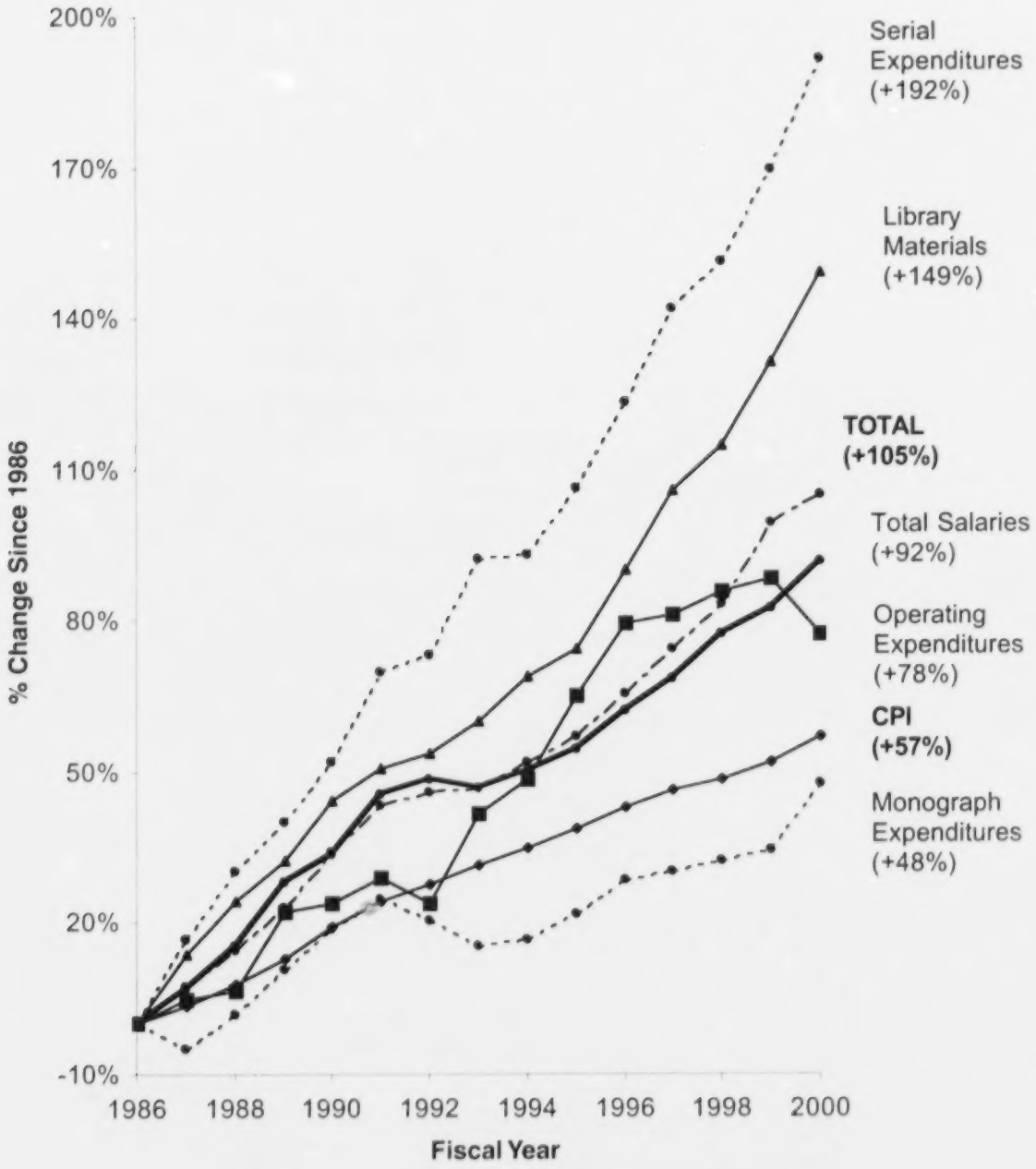
Emergency preparedness for libraries involves prevention, preparation, and response. Prevention is the reduction of risk, including inspecting, maintaining, and repairing library buildings. Preparation involves writing a disaster plan, assembling essential supplies, developing communications channels with essential service providers, and training staff to react effectively. Response is the skillful salvaging of collections damaged as the result of minor destructive incidents and catastrophic events.

STAFF AND USER EDUCATION: Protecting collections through teaching

The goal of preservation education programs is to build awareness within the library user community of the fragile, irreplaceable nature of research library collections, and to teach and encourage improved care and handling practices. Printed materials, training sessions, seminars, and conferences are typical communications vehicles employed for this purpose.

APPENDIX IV
Expenditure Trends in ARL Libraries, 1986-2000

(available at http://www.arl.org/stats/arlstat/graphs/2000t4.html)



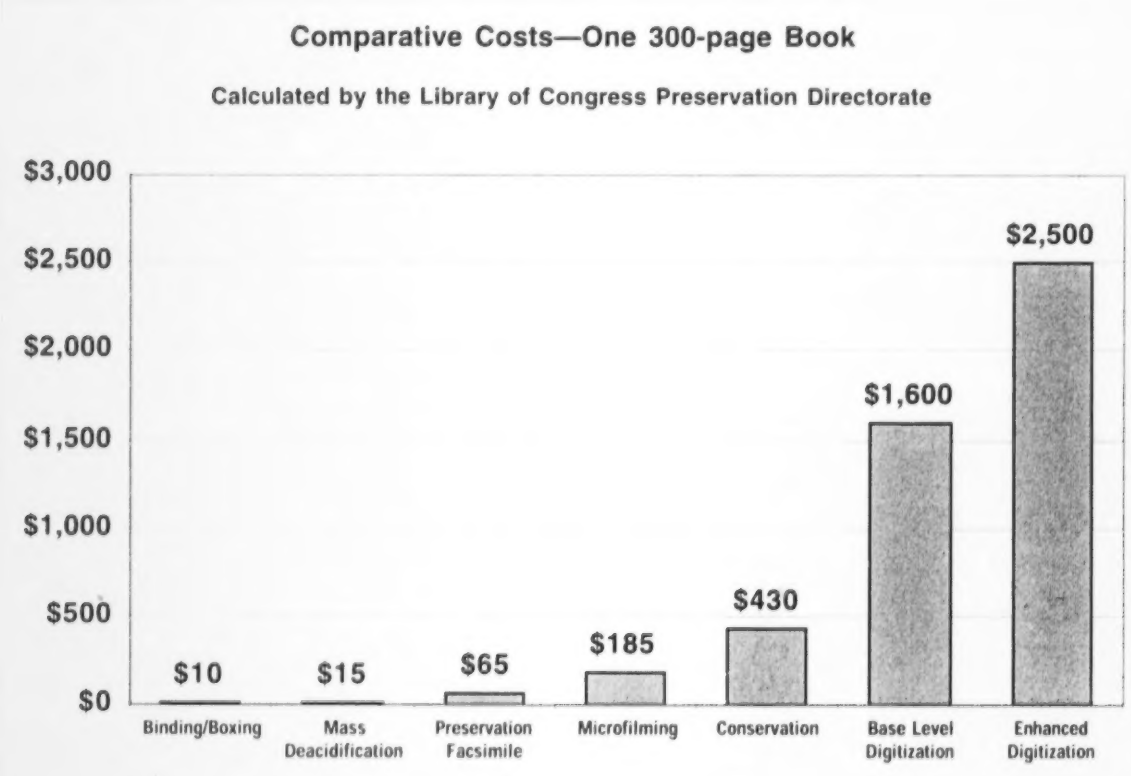
APPENDIX V
Non-print Holdings in ARL University Libraries

ARL university libraries report the following holdings as of June 30, 1999. Data are compiled from ARL Web site (http://fisher.lib.virginia.edu/newarl/).

Type of holding	# of libraries responding	# of holdings
Microform units	111	455,866,843
Government documents	97	41,827,671
Computer files	102	578,637
Archives and manus.	107	2,361,544
Cartographic materials	106	27,885,352
Graphic materials	101	64,083,533
Audio materials	104	5,385,634
Video and film	103	1,335,441

BEST COPY AVAILABLE

APPENDIX VI
Comparative Costs for Book Treatments



NOTES

Some materials that are microfilmed may subsequently be boxed. Objects that are bound may later be mass deacidified, and selective objects treated by a conservator this year may be digitized ten years hence. However, we do not generally think of various preservation and digitization options as being logically applied to the same materials in a library or an archives. By the same token, digitization and preservation activities cannot logically be substituted for one another. For example, digitization is not seen as a substitute for deacidification or conservation. The digitization technology, aimed at improving access to information, is not generally applied to the same materials that are saved in their original format through mass deacidification or conservation. Nor are the materials selected for deacidification the same materials that would normally be candidates for reformatting either through microfilming (due to advanced embrittlement of the paper) or through digitization (to promote much broader public access). Given these caveats, the relative costs of such activities continue to be of interest to persons engaged in managing library collections and archival materials and in making short-range and long-term plans

BEST COPY AVAILABLE

(notes, continued)

for preservation and access to media and the information it contains.

Costs

For Binding/Boxing and for Mass Deacidification, unit costs are rounded off to the nearest dollar; for each of the other activities in the chart, comparative unit costs are rounded up or down to the nearest increment of $5.00. For the first six functions, costs include contract, supply, staff and related administrative costs at LC. The second category for digitization, described below in the context of the *RLG DigiNews* article under "Enhanced Digitization," includes averaged costs for these activities at LC and the National Archives, as well as project costs from other digitization projects and from published sources.

Conservation

Hands-on conservation treatment, preventive conservation measures, and specialized housings are generally provided for materials that are both at risk and a high priority to an institution, as well as for rare and intrinsically valuable materials among permanent research collections. Given the wide variety of formats, media, and treatment challenges represented in rare books, it is facile to attempt to present a cost that purports to represent the "average" expenses incurred for professional conservation treatment of a single, representative book. Nevertheless, the $430-per-book "Conservation" cost given here represents the average, in-house, Library of Congress cost for treating rare books between 1998 and 2000 in the Book and Paper Section of the Preservation Directorate's Conservation Division. The costs, which generally ranged from $275 to $720 per volume during that period, were derived by dividing the total number of rare books treated into the aggregate costs for conservation materials, permanent and temporary staff salaries and benefits, and, when appropriate for special rare book treatment projects, the costs incurred by hiring additional conservators as part-time contractors to supplement LC staff resources.

Mass Deacidification

Deacidification is an economical approach to keeping books and unbound paper-based materials alive and available in usable form. The Library uses this technology to save important, endangered materials that are central to its mission, treating acidic and slightly brittle items from the general and special collections and the Law Library that must be preserved. Current LC per-book deacidification cost is $13.40; adding $1.85/book for selection, book charging, packing, transportation, quality control, and reshelving, the total cost is $15.25/book.

Preservation Facsimile

The exact reproduction of a printed volume, including illustrated matter. This was formerly called "Preservation Photocopying"; but the characterization has been changed because image capture is now usually accomplished by digital scanning of the original. Printouts on preservation-quality paper are bound according to library standards. The cost for a 300-page book represented here includes aggregated costs for scanning, covering material and binding, labeling, insertion of tattle (security) tag, and associated administrative expenses such as invoicing.

Digitization

Two approaches to digitization serve preservation and access goals. Base level digitization offers a method of reformatting at-risk materials like brittle books. Enhanced digitization offers improved access to materials that have high informational and/or visual value such as significant manuscript or graphic collections including, for example, folklore artifacts, grey-scale images, exhibit-quality photographs, pamphlets, broadsides, scripts, music scores, or the correspondence of famous officials, writers, composers, or scientists. The cost of enhanced digitization may be justified for special collection and non-book materials for which research demand is high.

Base Level Digitization: For digitization of at-

(notes, continued)

risk materials, the LC Preservation Directorate requires high-quality master image files that allow for a broad range of processing and display options; economical digitizing methods including both on-site and off-site capture of disbound materials; machine-readable, minimally-encoded text generated by fully automated processes of Optical Character Recognition and mark-up; and basic bibliographic description. On-site capture is essential for rare materials that must be dealt with in-house—i.e. materials that will not be sent off-site for reformatting. This model provides a baseline of digital images and text onto which "value-added" enhancements (such as the essays, finding aids, and the kinds of exhibits done for American Memory collections) can later be added. These minimum requirements, without "value-added" enhancements or elaborate bibliographic data, contribute to lower per-page costs. The average cost for digitizing a book page, including scanning, metadata creation, automated generation of OCR and minimally-encoded text, and associated activities, including identifying and preparing materials, quality control, and project management, is $5.32. For a brief, 300-page book, this works out to $1,600.00.

Enhanced Digitization: Costs represent an "adjusted average" derived from two rounds of LC/Ameritech projects, NARA's Electronic Access Project, and various other projects and published sources. (See Steven Puglia, "The Costs of Digital Imaging Projects," *RLG DigiNews*, Vol. 3, No. 5, pp.1-6, Oct. 15, 1999.) The cost is derived from a broad range of digitizing models, which include low-, medium-, and high-quality digital images; item-level bibliographic description for various types of materials; and "value-added" enhancements such as essays, finding aids, or exhibits. Costs of creating machine-readable texts with complex SGML-encoding are not included. The average cost for digitizing a book page, including scanning, metadata creation, creation of enhancements, and other associated activities such as identifying and preparing materials, quality control, and project management, is given as $8.35 in Table 2 of the *DigiNews* article referenced above. For a brief, 300-page book, this works out to $2,500.00.

As confirmation of this rough $2,500.00 cost estimate for digitizing a small book, see also: Mark Y. Herring, "10 Reasons Why the Internet Is No Substitute for a Library," in *American Libraries*, Vol. 32. No. 4 (April 2001), pp. 76-78. The author describes Questia Media, Inc.'s recent expenditure of $125 million, digitizing 50,000 books that are now available electronically. Again, this comes out to roughly $2,500.00 per book. Note: As with the "base line" and "enhanced digitization" costs given above, this *does not include the expense of maintaining digital versions over time.*

APPENDIX VII
National Recording Preservation Act of 2000

H. R. 4846

One Hundred Sixth Congress of the United States of America

AT THE SECOND SESSION

Begun and held at the City of Washington on Monday, the twenty-fourth day of January, two thousand

An Act

To establish the National Recording Registry in the Library of Congress to maintain and preserve sound recordings that are culturally, historically, or aesthetically significant, and for other purposes.

Be it enacted by the Senate and House of Representatives of the United States of America in Congress assembled,

SECTION 1. SHORT TITLE.

This Act may be cited as the "National Recording Preservation Act of 2000".

TITLE I—SOUND RECORDING PRESERVATION BY THE LIBRARY OF CONGRESS

Subtitle A—National Recording Registry

SEC. 101. NATIONAL RECORDING REGISTRY OF THE LIBRARY OF CONGRESS.

The Librarian of Congress shall establish the National Recording Registry for the purpose of maintaining and preserving sound recordings that are culturally, historically, or aesthetically significant.

SEC. 102. DUTIES OF LIBRARIAN OF CONGRESS.

(a) ESTABLISHMENT OF CRITERIA AND PROCEDURES.—For purposes of carrying out this subtitle, the Librarian shall—

(1) establish criteria and procedures under which sound recordings may be included in the National Recording Registry, except that no sound recording shall be eligible for inclusion in the National Recording Registry until 10 years after the recording's creation;

(2) establish procedures under which the general public may make recommendations to the National Recording Preservation Board established under subtitle C regarding the inclusion of sound recordings in the National Recording Registry; and

(3) determine which sound recordings satisfy the criteria established under paragraph (1) and select such recordings for inclusion in the National Recording Registry.

(b) PUBLICATION OF SOUND RECORDINGS IN THE REGISTRY.—The Librarian shall publish in the Federal Register the name of

BEST COPY AVAILABLE

each sound recording that is selected for inclusion in the National Recording Registry.

SEC. 103. SEAL OF THE NATIONAL RECORDING REGISTRY.

(a) IN GENERAL.—The Librarian shall provide a seal to indicate that a sound recording has been included in the National Recording Registry and is the Registry version of that recording.

(b) USE OF SEAL.—The Librarian shall establish guidelines for approval of the use of the seal provided under subsection (a), and shall include in the guidelines the following:

(1) The seal may only be used on recording copies of the Registry version of a sound recording.

(2) The seal may be used only after the Librarian has given approval to those persons seeking to apply the seal in accordance with the guidelines.

(3) In the case of copyrighted mass distributed, broadcast, or published works, only the copyright legal owner or an authorized licensee of that copyright owner may place or authorize the placement of the seal on any recording copy of the Registry version of any sound recording that is maintained in the National Recording Registry Collection in the Library of Congress.

(4) Anyone authorized to place the seal on any recording copy of any Registry version of a sound recording may accompany such seal with the following language: "This sound recording is selected for inclusion in the National Recording Registry by the Librarian of Congress in consultation with the National Recording Preservation Board of the Library of Congress because of its cultural, historical, or aesthetic significance."

(c) EFFECTIVE DATE OF THE SEAL.—The use of the seal provided under subsection (a) with respect to a sound recording shall be effective beginning on the date the Librarian publishes in the Federal Register (in accordance with section 102(b)) the name of the recording, as selected for inclusion in the National Recording Registry.

(d) PROHIBITED USES OF THE SEAL.—

(1) PROHIBITION ON DISTRIBUTION AND EXHIBITION.—No person may knowingly distribute or exhibit to the public a version of a sound recording or any copy of a sound recording which bears the seal described in subsection (a) if such recording—

(A) is not included in the National Recording Registry; or

(B) is included in the National Recording Registry but has not been approved for use of the seal by the Librarian pursuant to the guidelines established under subsection (b).

(2) PROHIBITION ON PROMOTION.—No person may knowingly use the seal described in subsection (a) to promote any version of a sound recording or recording copy other than a Registry version.

(e) REMEDIES FOR VIOLATIONS.—

(1) JURISDICTION.—The several district courts of the United States shall have jurisdiction, for cause shown, to prevent and restrain violations of subsection (d).

(2) RELIEF.—

BEST COPY AVAILABLE

(A) REMOVAL OF SEAL.—Except as provided in subparagraph (B), relief for violation of subsection (d) shall be limited to the removal of the seal from the sound recording involved in the violation.

(B) FINE AND INJUNCTIVE RELIEF.—In the case of a pattern or practice of the willful violation of subsection (d), the court may order a civil fine of not more than $10,000 and appropriate injunctive relief.

(3) LIMITATION OF REMEDIES.—The remedies provided in this subsection shall be the exclusive remedies under this title, or any other Federal or State law, regarding the use of the seal described in subsection (a).

SEC. 104. NATIONAL RECORDING REGISTRY COLLECTION OF THE LIBRARY OF CONGRESS.

(a) IN GENERAL.—All copies of sound recordings on the National Recording Registry that are received by the Librarian under subsection (b) shall be maintained in the Library of Congress and be known as the "National Recording Registry Collection of the Library of Congress". The Librarian shall by regulation and in accordance with title 17, United States Code, provide for reasonable access to the sound recordings and other materials in such collection for scholarly and research purposes.

(b) ACQUISITION OF QUALITY COPIES.—

(1) IN GENERAL.—The Librarian shall seek to obtain, by gift from the owner, a quality copy of the Registry version of each sound recording included in the National Recording Registry.

(2) LIMIT ON NUMBER OF COPIES.—Not more than one copy of the same version or take of any sound recording may be preserved in the National Recording Registry. Nothing in the preceding sentence may be construed to prohibit the Librarian from making or distributing copies of sound recordings included in the Registry for purposes of carrying out this Act.

(c) PROPERTY OF UNITED STATES.—All copies of sound recordings on the National Recording Registry that are received by the Librarian under subsection (b) shall become the property of the United States Government, subject to the provisions of title 17, United States Code.

Subtitle B—National Sound Recording Preservation Program

SEC. 111. ESTABLISHMENT OF PROGRAM BY LIBRARIAN OF CONGRESS.

(a) IN GENERAL.—The Librarian shall, after consultation with the National Recording Preservation Board established under subtitle C, implement a comprehensive national sound recording preservation program, in conjunction with other sound recording archivists, educators and historians, copyright owners, recording industry representatives, and others involved in activities related to sound recording preservation, and taking into account studies conducted by the Board.

(b) CONTENTS OF PROGRAM SPECIFIED.—The program established under subsection (a) shall—

BEST COPY AVAILABLE

(1) coordinate activities to assure that efforts of archivists and copyright owners, and others in the public and private sector, are effective and complementary;

(2) generate public awareness of and support for these activities;

(3) increase accessibility of sound recordings for educational purposes;

(4) undertake studies and investigations of sound recording preservation activities as needed, including the efficacy of new technologies, and recommend solutions to improve these practices; and

(5) utilize the audiovisual conservation center of the Library of Congress at Culpeper, Virginia, to ensure that preserved sound recordings included in the National Recording Registry are stored in a proper manner and disseminated to researchers, scholars, and the public as may be appropriate in accordance with title 17, United States Code, and the terms of any agreements between the Librarian and persons who hold copyrights to such recordings.

SEC. 112. PROMOTING ACCESSIBILITY AND PUBLIC AWARENESS OF SOUND RECORDINGS.

The Librarian shall carry out activities to make sound recordings included in the National Recording Registry more broadly accessible for research and educational purposes and to generate public awareness and support of the Registry and the comprehensive national sound recording preservation program established under this subtitle.

Subtitle C—National Recording Preservation Board

SEC. 121. ESTABLISHMENT.

The Librarian shall establish in the Library of Congress a National Recording Preservation Board whose members shall be selected in accordance with the procedures described in section 122.

SEC. 122. APPOINTMENT OF MEMBERS.

(a) SELECTIONS FROM LISTS SUBMITTED BY ORGANIZATIONS.—

(1) IN GENERAL.—The Librarian shall request each organization described in paragraph (2) to submit a list of three candidates qualified to serve as a member of the Board. The Librarian shall appoint one member from each such list, and shall designate from that list an alternate who may attend at Board expense those meetings which the individual appointed to the Board cannot attend.

(2) ORGANIZATIONS DESCRIBED.—The organizations described in this paragraph are as follows:

(A) National Academy of Recording Arts and Sciences (NARAS).

(B) Recording Industry Association of America (RIAA).

(C) Association for Recorded Sound Collections (ARSC).

(D) American Society of Composers, Authors and Publishers (ASCAP).

(E) Broadcast Music, Inc. (BMI).

BEST COPY AVAILABLE

(F) Songwriters Association (SESAC).
(G) American Federation of Musicians (AF of M).
(H) Music Library Association.
(I) American Musicological Society.
(J) National Archives and Record Administration.
(K) National Association of Recording Merchandisers (NARM).
(L) Society for Ethnomusicology.
(M) American Folklore Society.
(N) Country Music Foundation.
(O) Audio Engineering Society (AES).
(P) National Academy of Popular Music.
(Q) Digital Media Association (DiMA).

(b) OTHER MEMBERS.—In addition to the members appointed under subsection (a), the Librarian may appoint not more than five members-at-large. The Librarian shall select an alternate for each member-at-large, who may attend at Board expense those meetings that the member-at-large cannot attend.

(c) CHAIR.—The Librarian shall appoint one member of the Board to serve as Chair.

(d) TERM OF OFFICE.—

(1) TERMS.—The term of each member of the Board shall be 4 years, except that there shall be no limit to the number of terms that any individual member may serve.

(2) REMOVAL OF MEMBER OR ORGANIZATION.—The Librarian shall have the authority to remove any member of the Board (or, in the case of a member appointed under subsection (a)(1), the organization that such member represents) if the member or organization over any consecutive 2-year period fails to attend at least one regularly scheduled Board meeting.

(3) VACANCIES.—A vacancy in the Board shall be filled in the manner in which the original appointment was made under subsection (a), except that the Librarian may fill the vacancy from a list of candidates previously submitted by the organization or organizations involved. Any member appointed to fill a vacancy shall be appointed for the remainder of the term of the member's predecessor.

SEC. 123. SERVICE OF MEMBERS; MEETINGS.

(a) REIMBURSEMENT OF EXPENSES.—Members of the Board shall serve without pay, but may receive travel expenses, including per diem in lieu of subsistence, in accordance with sections 5702 and 5703 of title 5, United States Code.

(b) CONFLICT OF INTEREST.—The Librarian shall establish rules and procedures to address any potential conflict of interest between a member of the Board and responsibilities of the Board.

(c) MEETINGS.—The Board shall meet at least once each fiscal year. Meetings shall be at the call of the Librarian.

(d) QUORUM.—Eleven members of the Board shall constitute a quorum for the transaction of business.

SEC. 124. RESPONSIBILITIES OF BOARD.

(a) REVIEW AND RECOMMENDATION OF NOMINATIONS FOR NATIONAL RECORDING REGISTRY.—

(1) IN GENERAL.—The Board shall review nominations of sound recordings submitted to it for inclusion in the National Recording Registry and advise the Librarian, as provided in subtitle A, with respect to the inclusion of such recordings

BEST COPY AVAILABLE

in the Registry and the preservation of these and other sound recordings that are culturally, historically, or aesthetically significant.

(2) SOURCE OF NOMINATIONS.—The Board shall consider for inclusion in the National Recording Registry nominations submitted by the general public as well as representatives of sound recording archives and the sound recording industry (such as the guilds and societies representing sound recording artists) and other creative artists.

(b) STUDY AND REPORT ON SOUND RECORDING PRESERVATION AND RESTORATION.—The Board shall conduct a study and issue a report on the following issues:

(1) The current state of sound recording archiving, preservation and restoration activities.

(2) Taking into account the research and other activities carried out by or on behalf of the National Audio-Visual Conservation Center at Culpeper, Virginia—

(A) the methodology and standards needed to make the transition from analog "open reel" preservation of sound recordings to digital preservation of sound recordings; and

(B) standards for access to preserved sound recordings by researchers, educators, and other interested parties.

(3) The establishment of clear standards for copying old sound recordings (including equipment specifications and equalization guidelines).

(4) Current laws and restrictions regarding the use of archives of sound recordings, including recommendations for changes in such laws and restrictions to enable the Library of Congress and other nonprofit institutions in the field of sound recording preservation to make their collections available to researchers in a digital format.

(5) Copyright and other laws applicable to the preservation of sound recordings.

SEC. 125. GENERAL POWERS OF BOARD.

(a) IN GENERAL.—The Board may, for the purpose of carrying out its duties, hold such hearings, sit and act at such times and places, take such testimony, and receive such evidence, as the Librarian and the Board consider appropriate.

(b) SERVICE ON FOUNDATION.—Two sitting members of the Board shall be appointed by the Librarian and shall serve as members of the board of directors of the National Recording Preservation Foundation, in accordance with section 152403 of title 36, United States Code.

Subtitle D—General Provisions

SEC. 131. DEFINITIONS.

As used in this title:

(1) The term "Librarian" means the Librarian of Congress.

(2) The term "Board" means the National Recording Preservation Board.

(3) The term "sound recording" has the meaning given such term in section 101 of title 17, United States Code.

(4) The term "publication" has the meaning given such term in section 101 of title 17, United States Code.

BEST COPY AVAILABLE

(5) The term "Registry version" means, with respect to a sound recording, the version of a recording first published or offered for mass distribution whether as a publication or a broadcast, or as complete a version as bona fide preservation and restoration activities by the Librarian, an archivist other than the Librarian, or the copyright legal owner can compile in those cases where the original material has been irretrievably lost or the recording is unpublished.

SEC. 132. STAFF; EXPERTS AND CONSULTANTS.

(a) STAFF.—The Librarian may appoint and fix the pay of such personnel as the Librarian considers appropriate to carry out this title.

(b) EXPERTS AND CONSULTANTS.—The Librarian may, in carrying out this title, procure temporary and intermittent services under section 3109(b) of title 5, United States Code, but at rates for individuals not to exceed the daily equivalent of the maximum rate of basic pay payable for level 15 of the General Schedule. In no case may a member of the Board (including an alternate member) be paid as an expert or consultant under this section.

SEC. 133. AUTHORIZATION OF APPROPRIATIONS.

There are authorized to be appropriated to the Librarian for each of the first 7 fiscal years beginning on or after the date of the enactment of this Act such sums as may be necessary to carry out this title, except that the amount authorized for any fiscal year may not exceed $250,000.

TITLE II—NATIONAL RECORDING PRESERVATION FOUNDATION

SEC. 201. NATIONAL RECORDING PRESERVATION FOUNDATION.

(a) IN GENERAL.—Part B of subtitle II of title 36, United States Code, is amended by inserting after chapter 1523 the following:

"CHAPTER 1524—NATIONAL RECORDING PRESERVATION FOUNDATION

"Sec.
"152401. Organization.
"152402. Purposes.
"152403. Board of directors.
"152404. Officers and employees.
"152405. Powers.
"152406. Principal office.
"152407. Provision and acceptance of support by Librarian of Congress.
"152408. Service of process.
"152409. Civil action by Attorney General for equitable relief.
"152410. Immunity of United States Government.
"152411. Authorization of appropriations.
"152412. Annual report.

"§ 152401. Organization

"(a) FEDERAL CHARTER.—The National Recording Preservation Foundation (in this chapter, the 'corporation') is a federally chartered corporation.

"(b) NATURE OF CORPORATION.—The corporation is a charitable and nonprofit corporation and is not an agency or establishment of the United States Government.

BEST COPY AVAILABLE

"(c) PERPETUAL EXISTENCE.—Except as otherwise provided, the corporation has perpetual existence.

"§ 152402. Purposes

"The purposes of the corporation are to—

"(1) encourage, accept, and administer private gifts to promote and ensure the preservation and public accessibility of the nation's sound recording heritage held at the Library of Congress and other public and nonprofit archives throughout the United States; and

"(2) further the goals of the Library of Congress and the National Recording Preservation Board in connection with their activities under the National Recording Preservation Act of 2000.

"§ 152403. Board of directors

"(a) GENERAL.—The board of directors is the governing body of the corporation.

"(b) MEMBERS AND APPOINTMENT.—(1) The Librarian of Congress (hereafter in this chapter referred to as the 'Librarian') is an ex officio nonvoting member of the board. Not later than 90 days after the date of the enactment of this chapter, the Librarian shall appoint the directors to the board in accordance with paragraph (2).

"(2)(A) The board consists of nine directors.

"(B) Each director shall be a United States citizen.

"(C) At least six directors shall be knowledgeable or experienced in sound recording production, distribution, preservation, or restoration, including two who are sitting members of the National Recording Preservation Board. These six directors shall, to the extent practicable, represent diverse points of view from the sound recording community.

"(3) A director is not an employee of the Library of Congress and appointment to the board does not constitute appointment as an officer or employee of the United States Government for the purpose of any law of the United States.

"(4) The terms of office of the directors are 4 years. An individual may not serve more than two consecutive terms.

"(5) A vacancy on the board shall be filled in the manner in which the original appointment was made.

"(c) CHAIR.—The Librarian shall appoint one of the directors as the initial chair of the board for a 2-year term. Thereafter, the chair shall be appointed and removed in accordance with the bylaws of the corporation.

"(d) QUORUM.—The number of directors constituting a quorum of the board shall be established under the bylaws of the corporation.

"(e) MEETINGS.—The board shall meet at the call of the Librarian for regularly scheduled meetings.

"(f) REIMBURSEMENT OF EXPENSES.—Directors shall serve without compensation but may receive travel expenses, including per diem in lieu of subsistence, in accordance with sections 5702 and 5703 of title 5.

"(g) LIABILITY OF DIRECTORS.—Directors are not personally liable, except for gross negligence.

BEST COPY AVAILABLE

"§ 152404. Officers and employees

"(a) SECRETARY OF THE BOARD.—(1) The Librarian shall appoint a Secretary of the Board to serve as executive director of the corporation. The Librarian may remove the Secretary.

"(2) The Secretary shall be knowledgeable and experienced in matters relating to—

"(A) sound recording preservation and restoration activities;

"(B) financial management; and

"(C) fundraising.

"(b) APPOINTMENT OF OFFICERS.—Except as provided in subsection (a) of this section, the board of directors appoints, removes, and replaces officers of the corporation.

"(c) APPOINTMENT OF EMPLOYEES.—Except as provided in subsection (a) of this section, the Secretary appoints, removes, and replaces employees of the corporation.

"(d) STATUS AND COMPENSATION OF EMPLOYEES.—Employees of the corporation (including the Secretary)—

"(1) are not employees of the Library of Congress;

"(2) shall be appointed and removed without regard to the provisions of title 5 governing appointments in the competitive service; and

"(3) may be paid without regard to chapter 51 and subchapter III of chapter 53 of title 5, except that an employee may not be paid more than the annual rate of basic pay for level 15 of the General Schedule under section 5107 of title 5.

"§ 152405. Powers

"(a) GENERAL.—The corporation may—

"(1) adopt a constitution and bylaws;

"(2) adopt a seal which shall be judicially noticed; and

"(3) do any other act necessary to carry out this chapter.

"(b) POWERS AS TRUSTEE.—To carry out its purposes, the corporation has the usual powers of a corporation acting as a trustee in the District of Columbia, including the power—

"(1) to accept, receive, solicit, hold, administer, and use any gift, devise, or bequest, either absolutely or in trust, of property or any income from or other interest in property;

"(2) to acquire property or an interest in property by purchase or exchange;

"(3) unless otherwise required by an instrument of transfer, to sell, donate, lease, invest, or otherwise dispose of any property or income from property;

"(4) to borrow money and issue instruments of indebtedness;

"(5) to make contracts and other arrangements with public agencies and private organizations and persons and to make payments necessary to carry out its functions;

"(6) to sue and be sued; and

"(7) to do any other act necessary and proper to carry out the purposes of the corporation.

"(c) ENCUMBERED OR RESTRICTED GIFTS.—A gift, devise, or bequest may be accepted by the corporation even though it is encumbered, restricted, or subject to beneficial interests of private persons, if any current or future interest is for the benefit of the corporation.

BEST COPY AVAILABLE

"§ 152406. Principal office

"The principal office of the corporation shall be in the District of Columbia. However, the corporation may conduct business throughout the States, territories, and possessions of the United States.

"§ 152407. Provision and acceptance of support by Librarian of Congress

"(a) PROVISION BY LIBRARIAN.—(1) The Librarian may provide personnel, facilities, and other administrative services to the corporation. Administrative services may include reimbursement of expenses under section 152403(f).

"(2) The corporation shall reimburse the Librarian for support provided under paragraph (1) of this subsection. Amounts reimbursed shall be deposited in the Treasury to the credit of the appropriations then current and chargeable for the cost of providing the support.

"(b) ACCEPTANCE BY LIBRARIAN.—The Librarian may accept, without regard to chapters 33 and 51 and subchapter III of chapter 53 of title 5 and related regulations, the services of the corporation and its directors, officers, and employees as volunteers in performing functions authorized under this chapter, without compensation from the Library of Congress.

"§ 152408. Service of process

"The corporation shall have a designated agent to receive service of process for the corporation. Notice to or service on the agent, or mailed to the business address of the agent, is notice to or service on the corporation.

"§ 152409. Civil action by Attorney General for equitable relief

"The Attorney General may bring a civil action in the United States District Court for the District of Columbia for appropriate equitable relief of the corporation—

"(1) engages or threatens to engage in any act, practice, or policy that is inconsistent with the purposes in section 152402 of this title; or

"(2) refuses, fails, or neglects to carry out its obligations under this chapter or threatens to do so.

"§ 152410. Immunity of United States Government

"The United States Government is not liable for any debts, defaults, acts, or omissions of the corporation. The full faith and credit of the Government does not extend to any obligation of the corporation.

"§ 152411. Authorization of appropriations

"(a) AUTHORIZATION.—There are authorized to be appropriated to the corporation for each of the first 7 fiscal years beginning on or after the date of the enactment of this chapter an amount not to exceed the amount of private contributions (whether in currency, services, or property) made to the corporation by private persons and State and local governments.

"(b) LIMITATION RELATED TO ADMINISTRATIVE EXPENSES.—Except as permitted under section 152407, amounts authorized

 BEST COPY AVAILABLE

under this section may not be used by the corporation for administrative expenses of the corporation, including salaries, travel, transportation, and overhead expenses.

"§ 152412. Annual report

"As soon as practicable after the end of each fiscal year, the corporation shall submit a report to the Librarian for transmission to Congress on the activities of the corporation during the prior fiscal year, including a complete statement of its receipts, expenditures, and investments."

(b) CLERICAL AMENDMENT.—The table of chapters for part II of subtitle II of title 36, United States Code, is amended by inserting after the item relating to chapter 1523 the following new item:

"1524. National Recording Preservation Foundation ..... 152401".

Speaker of the House of Representatives.

Vice President of the United States and
President of the Senate.

BEST COPY AVAILABLE